4-17-1993

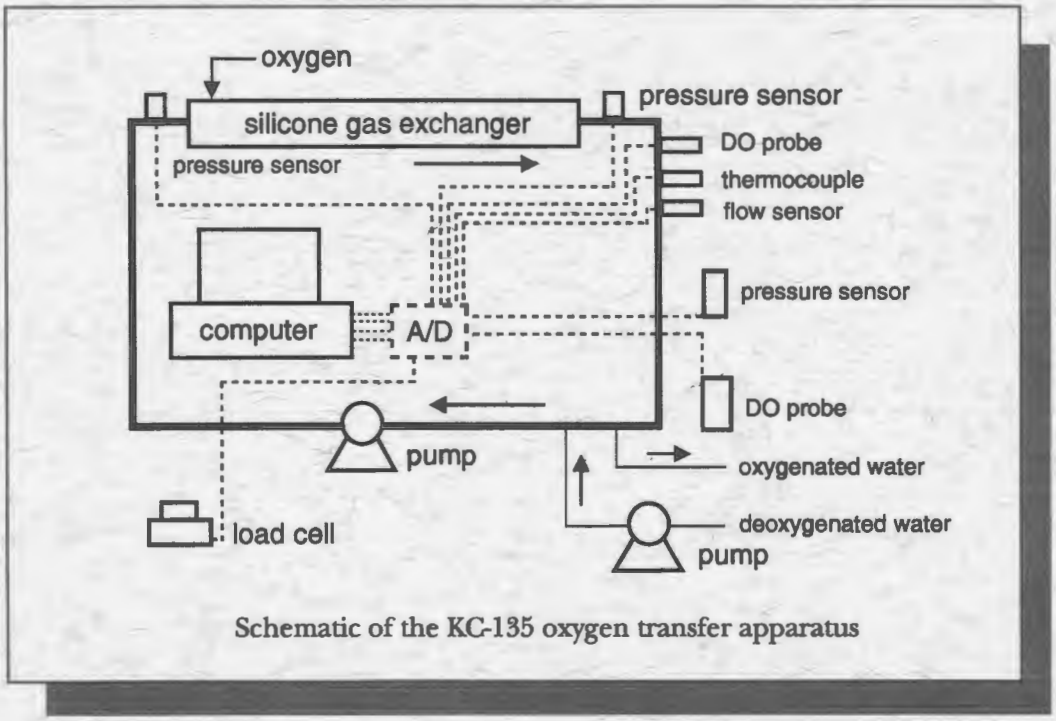
Proceedings of the 23rd Annual Biochemical Engineering Symposium

Roger G. Harrison
University of Oklahoma

Proceedings of the 23rd

Annual Biochemical Engineering Symposium

April 17, 1993, Norman, Oklahoma



Schematic of the KC-135 oxygen transfer apparatus

Roger G. Harrison
Editor

University of Oklahoma
School of Chemical Engineering and Materials Science

Proceedings of the Twenty-Third Annual Biochemical Engineering Symposium, April 17, 1993

Roger G. Harrison
Editor

School of Chemical Engineering and Materials Science
University of Oklahoma
Norman, Oklahoma 73019

Preface

The 23rd Annual Biochemical Engineering Syposium was held at the University of Oklahoma on April 17, 1993. The objectives of the symposium were to provide 1) a forum for informal discussion of biochemical engineering research being carried at the participating universities and 2) an opportunity for students to present and publish their work.

Thirteen papers presented at the symposium are included in the proceedings. Because final publication usually takes place in refereed journals, the articles included here are typically brief and often cover work in progress. The program of the symposium and a list of participants are included in the proceedings.

I would like to thank all those who participated in this symposium, especially those who presented papers or posters. Financial support for this symposium from the College of Engineering and the School of Chemical Engineering and Materials Science at the University of Oklahoma is gratefully acknowledged.

Roger G. Harrison
Editor

December 31, 1993

TABLE OF CONTENTS

23rd Annual Biochemical Engineering Symposium

University of Oklahoma
Norman, OK
April 17, 1993

Session 1: Oral Presentations (A235 Energy Center)

8:30 Welcome (Roger Harrison)

8:35 A Bioreactor Strategy for the Low Cost Production of RNA. Tony Marble, Eleni Chrisikos, and Robert Davis, University of Colorado-Boulder.

9:00 Development of a CELSS Bioreactor: Oxygen Transfer and Micromixing in Parabolic Flight. K.S. Wenger, P.E. Villeneuve, B.G. Thompson, T. Kedar, and E.H. Dunlop, Colorado State University.

9:25 A Three-Dimensional Culture System for Murine Bone Marrow Cells. John Highfill, Paul Todd, and Dhinakar Kompala, University of Colorado-Boulder.

9:50 Modeling and Estimation of States of Recombinant Fermentations Using Nonlinear Input/Output Models. Victor Saucedo and Nazmul Karim, Colorado State University.

10:15 Break

Session 2: (10:30 - Noon) Poster Presentations (Atrium, First Floor, Energy Center)

Analysis of Foreign Protein Overproduction in Recombinant CHO Cells: Effect of Growth Kinetics and Cell Cycle Traverse. Man Bock Gu, Paul Todd, and Dhinakar Kompala, University of Colorado-Boulder.

Particle Partitioning in Aqueous Biphasic Systems. Debra Hawker, Robert Davis, and Paul Todd, University of Colorado-Boulder; Gote Johansson, University of Lund.

Deadend Microfiltration of Bovine Serum Albumin Suspension Through Yeast Cake Layers and Asymmetric Polymeric Membranes. Naveen Arora and Robert Davis, University of Colorado-Boulder.

Viable Cell Recycle with an Inclined Settler in the Perfusion Culture of Suspended Recombinant Mammalian Cells. James Searles, Paul Todd, and Dhinakar Kompala, University of Colorado-Boulder.

Rotary and Tangential Crossflow Microfiltration for Protein Recovery from Cell Debris. Charles Parnham and Robert Davis, University of Colorado-Boulder.

Optimization of Crossflow Filtration Process Using Backpulsing as a Cleaning Method. Sanjeev Redkar, Robert Davis, and Fred Ramirez, University of Colorado-Boulder.

Flux of Sugars from Sugar Beets. Daryl Pulver and James Linden, Colorado State University.

Biodegradation of TNT by *Pseudomonas fluoresens* under Nitrate Respiration and Nitrogen Limiting Conditions. Patrick Gilcrease and Vincent Murphy, Colorado State University.

Preferential Degradation of Lignin in a Mixture of Water Hyacinth and Barley Straw by *Pleurotus ostreatus*. Pronab Das, Linda Henk, and Nazmul Karim, Colorado State University.

Enhancement of Polyelectrolyte Precipitation through the Genetic Fusion of Charged Peptide Tails to Proteins. Mark Niederauer and Charles Glatz, Iowa State University.

Stagewise Design for Concentration of Spirulina Suspensions Based on Lateral Migration with Flow-Thru Vertical Tees. N.M. Poflee, A.L. Rakow, M. Werntz, and M.N. Pons, Colorado State University.

Isolation and Characterization of Thermosensitive Mutants of Glucoamylase. Nathalie Flory, Maureen Gorman, Peter Reilly, and Clark Ford, Iowa State University; Birte Sevnsson, Carlsberg Laboratory, Copenhagen Valby, Denmark.

Food Grade Oxygen Reducing Membrane Bound Enzyme Systems. K. Tuitemwong, D.Y.C. Fung, and P. Tuitemwong, Kansas State University.

Monitoring the Fate of Toluene and Phenol in the Rhizoshpere. N. Muralidharan, Lawrence Davis, and Larry Erickson, Kansas State University.

Hydrodynamic Motions Associated with Bubble Coalescence and Breakup. T.Y. Yiin, L.A. Glasgow, and L.E. Erickson, Kansas State University.

A Method to Insert a DNA Fragment into a Double-Stranded Plasmid. Chris Haught, David Wilkinson, Kostas Zgafas, and Roger Harrison, University of Oklahoma.

12:00 Lunch

1:30 Tours/Demonstrations

Session 3: Oral Presentations (A235 Energy Center)

2:00 Expression and Purification of α-Human Atrial Natriuretic Peptide in *Escherichia coli* by Fusion with L-Asparaginase. Nien-Tung Ma and Roger Harrison, University of Oklahoma.

2:25 High Pressure Crystallization of Proteins. Mungara Saikumar, Charles Glatz, and Maurice Larson, Iowa State University.

2:50 Structure/Function Relationships in the Catalytic and Starch Binding Domains of Glucoamylase. Pedro Coutinho and Peter Reilly, Iowa State University.

3:15 Break

Session 4: Oral Presentations (A235 Energy Center)

3:30 Effects of Surfactant F68 and Well-Defined Shear on Membrane Lipids of SF9 Insect Cells. P. Yeh, G.A. Weisman, G.Y. Sun, and R.K. Bajpai, University of Missouri-Columbia.

3:55 A Novel Approach to Understanding the Antimicrobial Activity of Peptides. Naveen Pathak, Marie-Helene Janna, Gael Ruche, David McCarthy, and Roger Harrison, University of Oklahoma.

4:20 Mass Transfer in the Bioremediation of Soils Contaminated with Trapped Non-Aqueous Phase Liquids. Xiaoqing Yang, Larry Erickson, and L.T. Fan, Kansas State University.

4:45 Closing Comments (Roger Harrison)

LIST OF PARTICIPANTS

Colorado State University

Professor Vince Murphy, Pronab Das, Pat Gilcrease, Narendra M. Poflee, William Nagle, Daryl Pulver, Victor M. Saucedo, Kevin Wenger, and Sun Xinzhi

Iowa State University

Professor Chuck Glatz, Professor Peter Reilly, Scott Blonigen, Pedro Coutinho, Bipin Dalmia, Michael K. Dowd, Nathalie Flory, Zhong Gu, Meng H. Heng, Mark Niederauer, M.V. Saikumar, Mungara Saikumar, and Maribel Rodriguez Torres

Kansas State University

Professor Larry E. Erickson, Chris Banner, Prashant Gandhi, Muralidharan Narayanan, Satishkumar Santharam, Kooranee Tuitemwong, Pravate Tuitemwong, Tian-Yih Yiin, and Xiaoqing Yang,

University of Colorado

Professor Robert Davis, Professor Dhinakar S. Kompala, Naveen Arora, Tanya Chavez-Cropp, Debra Hawker, John Highfill, Ken Jones, Tony Marble, Scott Nichols, Charles Parnham, Sanjeev Redkar, Chris Roenfeldt, Jim Searles, Eve Tracey, and Barry Vant-Hull,

University of Missouri

Professor Rakesh Bajpai and Paul Yeh

University of Oklahoma

Professor Roger Harrison, Marie-Helene Janna, Jing Jing Lee, Nien-Tung Ma, Naveen Pathak, Choon Teo, and David Wilkinson

A LOW-COST BIOREACTOR STRATEGY FOR RNA SYNTHESIS

H. ANTHONY MARBLE*, ELENI CHRISIKOS, AND ROBERT H. DAVIS
*DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY
DEPARTMENT OF CHEMICAL ENGINEERING
UNIVERSITY OF COLORADO
BOULDER, COLORADO 80309-0424

ABSTRACT

The increased use of ribonucleic acid (RNA) in basic biological research and *in vitro* diagnostic applications has prompted the demand for large-scale, cost-effective RNA production strategies. RNA is currently made by either direct chemical or enzymatic synthesis. Scale-up is prohibitive due to the high material costs coupled with the limitations of batch-mode processes. Our approach is to harness the transcriptional activity of purified bacteriophage T7 RNA polymerase using solid-phase DNA templates in a continuous-flow, stirred-cell bioreactor. A direct comparison of activity reveals that immobilized DNA templates are transcribed at about one-half the initial rate of solution phase templates in conventional batch reactions. Despite this reduced transcription rate, equivalent yields of RNA are obtained if the reaction time period is extended. Immobilized templates store stably for periods of several months and are easily recovered from batch reactions for subsequent reuse. These findings indicate that solid-phase templates are ideal for continuous-flow and semi-batch applications, particularly since they are selectively retained in the bioreactor and thereby further facilitate downstream purification. Finally, we demonstrate this process by transcribing RNA from immobilized DNA templates in a stirred-cell reactor operated in a semi-batch mode.

INTRODUCTION

Ribonucleic acid is a ubiquitous biopolymer integral to protein synthesis and cellular metabolism. It is generally composed of four basic monomers interlinked with phosphate diester linkages to form a directional (5'-3') polymer. In nature RNA exists in various lengths and shapes, ranging from short oligoribonucleotide primers for DNA synthesis to the very large ribosomal RNAs which form part of the translational machinery active during protein synthesis. The purine and pyrimidine bases of RNA interact via ring-stacking and hydrogen bonding networks to fold the polymer chains into compact structures containing surfaces capable of providing specific binding sites for a variety of ligands, including proteins, nucleic acids, divalent metal ions, and carbohydrates.[1] The 2'-hydroxyl group provides additional hydrogen bonding capacity and serves as a modest nucleophile in the presence of electropositive magnesium ions. In fact, some RNAs are capable of topologically juxtaposing surface recognition and catalysis elements to promote metal cofactor driven chemical reactions such as transesterification and phosphodiester bond hydrolysis.[2] These sophisticated attributes have prompted investigators to pursure therapeutic applications based on RNA catalysis and specific ligand binding activities.

RNA is currently produced by either direct chemical or enzymatic synthesis at substantial cost. The former is assembled from the stepwise addition of fully protected, activated monomers in free solution (or on a solid support matrix) followed

by chemical deprotection. The latter is obtained from cell-free, *in vitro* transcription using purified bacteriophage RNA polymerase, either plasmid or synthetic DNA templates, and ribonucleoside triphosphate precursors. Since both processes operate in a batch mode, massive production is limited by economy of scale. Chemical synthesis is favored for creation of short RNA species (less than twenty nucleotides in length), whereas enzymatic synthesis is preferred for larger molecules (ranging from twenty to several hundred bases in length). Enzymatic approaches appear more versatile since production costs do not appreciate linearly with increasing polymer length. In contrast, chemical synthesis yields decrease with increasing length due to the exponential effect of inefficient stepwise addition reactions.

Bacteriophage T7 RNA polymerase is a 98.5 KD, single polypeptide chain containing a magnesium-requiring, DNA-dependent, RNA synthesis activity.[3] The enzyme specifically recognizes and binds to an eighteen base-paired double-stranded promoter region required for transcription initiation. The polymerase synthesizes RNA in the 5'-3' direction, consuming ribonucleoside triphosphate (NTP) precursors using DNA as a template. The nascent transcript forms a transient RNA\DNA hybrid mediated by Watson-Crick base pairs which continually form at its growing 3'-end. The enzyme displays temperature and pH optima at 37°C, pH 8.3 and requires a reducing agent to maintain one or more free (reduced) cysteine residues which are essential for transcriptional activity.[4] The enzyme has been cloned into and overexpressed in *E. coli*, from which it is subsequently purified to near homogeneity for cell-free, *in vitro* transcription applications.[5]

T7 RNA polymerase is active on double-stranded, T7 promoter-containing plasmid templates as well as synthetic DNA templates (composed of a double-stranded promoter region directly upstream of a single-stranded coding region).[6] The former are favored for synthesis of large RNA molecules (greater than fifty bases in length), whereas the latter are used for smaller RNA transcripts. Run-off transcription from linearized plasmid templates proceeds with high fidelity, yielding one predominant RNA product species. In contrast, enzyme mediated synthesis from synthetic DNA templates is not as processive, yielding a distribution of products varying in size from short, premature, (abortive) transcripts to high molecular weight run-on (extended) RNAs. This complication requires downstream purification (denaturing polyacrylamide gel electrophoresis or high performance liquid chromatography) to afford the desired product devoid of contaminating RNA or DNA. Abortive transcripts arise from premature dissociation of RNA polymerase (or the nascent RNA transcript) from the active ternary complex. Run-on products presumably arise from template independent polymerization events which are believed to be promoted by the unusually high enzyme and substrate concentrations. In theory, one way to minimize abortive transcripts is to stabilize the active ternary complex from premature dissociation. Polyethylene glycol (MW 8,000) has been reported to stimulate transcription yields[7], presumably by lowering the chemical potential of the solvent species and thereby creating excluded volume effects which are thought to stabilize the ternary complex.

A material cost analysis reveals that the purified RNA polymerase is the most costly reaction component, followed by the DNA templates and NTPs, respectively. The cost of the remaining materials (transcription buffer components) are trivial in comparison. A logical approach is to maximize the transcriptional productivity, in terms of the number of RNAs transcribed per polymerase molecule, thereby

reducing enzyme costs. Furthermore, since most protocols discard synthetic DNA templates (or remove them by enzymatic treatment) following a single batch reaction, material costs are further reduced by recycling templates in multiple rounds of reactions. By coupling these two strategies, transcriptional productivity is increased with respect to both enzyme and template, while material costs are lowered.

Our approach is to exploit the use of immobilized synthetic DNA templates in a stirred-cell, transcription reactor for the production of short RNA species (twenty to fifty nucleotides in length). The bioreactor is configured such that reactants are delivered via a feed line to the reactor core and products are removed by passing through an ultrafiltration membrane (nominal molecular weight cut-off 100,000 daltons) to the exit stream. The ultrafiltration membrane is composed of a low protein-binding polymer matrix and serves to selectively retain both the RNA polymerase and the immobilized DNA templates in the reactor core. The transcription reactor can be operated in either semi-batch or continuous-flow modes, depending on user preferences. In practice, the product size is limited to about 100 nucleotides, since RNAs of greater molecular weight are not expected to cross the ultrafiltration membrane into the exit line.

MATERIALS AND METHODS

DNA TEMPLATE IMMOBILIZATION

Synthetic DNA molecules were prepared by Genosys Biotechnologies Inc. (The Woodlands, TX), using conventional phosphoramidite chemistry. The transcriptional template codes for a 28-nucleotide pseudo-knot RNA molecule which specifically binds to the human immunodeficiency virus reverse transcriptase (HIV RT) primer binding site. The template consists of two DNA molecules (36 and 61 nucleotides, respectively) which hybridize to form a unique thirty three base pair hybrid, containing a class II T7 consensus promoter sequence upstream of a single-stranded RNA coding region. The top strand (36-mer) is modified with a 5'-terminal biotin group which functions to immobilize the oligomer to a solid support matrix which contains streptavidin protein, while the RNA coding (bottom) strand (61-mer) remains unmodified. The template was designed such that the 5'-biotin moiety is followed by a single-stranded trinucleotide overhang which acts as a linker arm to provide adequate spatial separation of the template from the support. The equilibrium dissociation constant of the biotin-streptavidin complex is extremely low (10^{-15} M), resulting in high affinity binding, despite the noncovalent nature of the interaction.

Immobilized templates were prepared by adding 7,500 picomoles of purified 5'-biotin DNA (36-mer) to 1.0 ml of a 50% (v/v) solution containing streptavidin-coated agarose beads in 0.02% sodium azide (Pierce Immunochemicals, Rockford, IL) and incubating overnight at 4°C with gentle mixing. The supernatant was removed and the beads were washed successively (4 x 0.5 ml) with sterile, deionized water. The unbound DNA in the original supernatant and the subsequent washes were quantified spectrophotometrically (260 nm or λ_{max}) to determine the amount of 5'-biotin DNA remaining on support. Typically, 6,000 to 7,000 picomoles of DNA remain bound, corresponding to a DNA concentration of 12 to 14 μM on the surface of the beads. Since the support contains 1.5 mg of streptavidin/ml of beads, the maximal theoretical biotin binding capacity is calculated to be 50 μM (based on 4:1 biotin/streptavidin binding stoichiometry).

The immobilized top strand is then incubated (overnight at 4°C) with a 1.5 molar excess of the bottom RNA coding strand (61-mer) to form the active template. Following DNA hybridization, the support is washed (3 x 0.5 ml sterile, deionized water) and quantified for the presence of unbound DNA. Solid phase templates are stored at 4°C in 0.02% sodium azide (preservative) until further use. In practice, we find that immobilized DNA templates are stable for periods of several months without detectable loss of transcriptional activity.

BATCH MODE TRANSCRIPTION REACTIONS

Transcription was assayed by the incorporation of radioactively labeled uridine triphosphate ([α-^{32}P]-UTP, 800 Curies/millimole, DuPont/NEN; Wilmington, DE) into RNA products at physiological temperature. Small-scale (100 µl) transcription reactions were performed using either solution or solid phase DNA templates and standard reaction conditions (described below), except that polyethylene glycol was omitted from all reactions. Timepoints (5 µl reaction aliquots) were removed at successive intervals, quenched on ice with 75% deionized formamide and loaded onto 20% (19:1 crosslinking) denaturing (8M urea) polyacrylamide gels for electrophoresis (600 Volts, 15 milliamps, for 3 hours). The gel-fractionated RNA products were visualized and the radioactivity was quantified using a phosphorescent imager system (Molecular Dynamics; Sunnyvale, CA).

Transcription Reaction Mixture:

40 mM Tris-HCl (pH 8.1)	1 mM Adenosine Triphosphate
20 mM Magnesium Chloride	1 mM Cytosine Triphosphate
5 mM Dithiothreitol	1 mM Guanosine Triphosphate
1 mM Spermidine-HCl	0.1 mM Uridine Triphosphate
0.01 % Triton X-100	0.125 µM [α-^{32}P]-UTP (10 µCi)
1 µM DNA Template	0.023 mg/ml T7 RNA Polymerase

SEMI-BATCH, STIRRED-CELL TRANSCRIPTION REACTIONS

Preparative scale (1.0 ml) transcription reactions were performed in a stirred-cell reactor (Amicon; Beverly, MA) operated in semi-batch mode, using either solution phase or immobilized DNA templates (1 µM reaction concentration). The stir-cell was equipped with a low protein-binding, ultrafiltration membrane (Amicon YM100) of nominal molecular weight cut-off 100,000 daltons. The transcription reaction conditions were identical to those listed above, except that the concentration of each NTP was 1 mM and no radioactivity was used. The stir-cell was loaded with a 1.0 ml reaction mixture and T7 RNA polymerase (0.023 mg/ml in reaction) was added at time zero. The reaction was allowed to incubate (37°C) for one hour on a magnetic stirrer (low spin setting), prior to the removal of the first aliquot (0.5 ml) from the exit line. At this time, the reaction was supplemented with additional transcription buffer containing 1 mM NTPs (0.5 ml), and the reaction was allowed to proceed for another hour. This semi-batch reaction cycle was performed for eight consecutive hours, with 0.5 ml aliquots removed from the exit line every hour. In some cases, fresh enzyme (0.023 mg/ml) was added along with the buffer and NTPs midway through the reaction (4 hour timepoint). Transcriptional activity was assayed qualitatively by fractionating RNA products on 20% (19:1 crosslinking) denaturing (8M urea) polyacrylamide gels, and staining overnight with stains-all dye (Sigma Chemical Co.; St. Louis, MO). Following destaining, transcriptional activity is assessed by the intensity of the dye present in each lane of the gel (RNA stains a blue color).

RESULTS AND DISCUSSION

IMMOBILIZED VS. SOLUTION PHASE DNA TEMPLATES

A direct comparison of the transcriptional activities of solid and solution phase templates in batch reactions (100 µl volume) is depicted in figure 1. The graph measures the cumulative [α-^{32}P]-UTP incorporation as a function of time. The data consist of numerical averages obtained from two separate experiments. This plot reveals that immobilized DNA templates sustain transcription at about one-half the initial rate of their solution phase counterparts. Despite the reduced transcriptional rate, equivalent yields of RNA are obtained if the reaction time period is doubled from two to four hours. In each case the reaction conditions were identical (1 µM DNA, 0.023 mg/ml T7 RNA polymerase), including the templates which differed only by the presence or absence of a 5'-terminal biotin moiety.

A lane by lane comparison of transcription from solid (I) and solution phase (S) templates is presented in the phosphorescent gel image displayed in figure 2. This reaction sequence is composed from individual timepoints obtained during the first hour of the transcription. Electrophoresis separates molecules on the basis of their charge to mass ratio. Hence, unincorporated UTP accumulates at the bottom of each lane followed by RNA products of increasing nucleotide length. Interestingly, a markedly different product distribution is obtained from immobilized compared to solution phase templates. The pattern of abortive (premature) transcripts obtained from solid phase templates is shifted down in size relative to solution. Furthermore, solution phase templates appear to generate more of the undesired run-on (extended) transcripts than their immobilized counterparts. The full-length RNA product is one of the few bands in common between the solution and solid phase reactions and is the most prominent RNA species in the reaction at sixty minutes (band position marked by the arrow in figure 2).

We interpret that, to a first approximation, the reduced transcriptional rate observed with immobilized templates arises from a decrease in the rate of transcription initiation (the rate limiting step of the reaction) rather than a reduced polymerization rate. Binding of the polymerase to the promoter is certainly influenced by the sterics of the surrounding environment. Despite the presence of the spacer region on the immobilized template, promoter binding may not be optimal due to overall steric constraints of the solid support matrix. We disfavor the possibility that the reduced transcriptional rate reflects a decrease in polymerization rate for several reasons. A decrease in the polymerization rate should lead to an increase in the prevalence of abortive transcripts, since longer polymerization times should promote dissociation of the active ternary complex. Since no general increase in abortive products is observed, a decreased polymerization rate is unlikely. Secondly, T7 RNA polymerase is a highly processive enzyme capable of incorporating 200-300 nucleotiudes per second[8]. Consequently with short templates, the overall rate of reaction is limited primarily by the initiation step. Since transcription intiation is rate limiting, there is no direct reason to assume a difference in polymerization rates between solid and solution phase templates.

The observation that solid phase templates yield lower amounts of undesired run-on transcripts is fortuitous and completely unexpected. At present we cannot provide an adequate explanation for this phenomenon. This is partly because the mechanics of run-on product formation is poorly understood. Nonetheless, this

finding is encouraging for several reasons. The absence of contaminating run-on transcripts directly facilitates downstream purification and product recovery from electrophoresis gels or HPLC columns. Secondly, the reduction in run-on transcripts minimizes unwanted reaction byproducts, which ultimately increases the transcriptional efficiency of the desired RNA product species. Finally, the reduction in run-on transcripts implies a concomitant reduction in the release of inorganic pyrophoshate (released in every nucleotidyl-transfer reaction) associated with run-on RNA synthesis. Since inorganic pyrophosphate is a known inhibitor of T7 RNA polymerase activity, reaction efficiency can be improved by limiting its accumulation.

Another advantage of immobilized templates is that they can be easily recovered from the crude reaction mixture for reuse in subsequent rounds of transcription reactions. Our results indicate that solid phase DNA templates can be successively reused up to four times without detectable loss in activity, despite being washed (2 x 0.5 ml) with water between rounds of reaction. Continued use beyond four rounds ultimately results in decreased yields, presumably due to template instability. Since the template is composed of two separate strands held together by noncovalent hydrogen-bonding interactions, template dissociation is a realistic possibility. Although template instability is the most likely cause of the reduced yields obtained in successive reactions, at this time we cannot altogether rule out other factors. Nevertheless, template reuse minimizes material costs and increases transcriptional productivity in terms of the number of transcripts generated per individual template molecule.

TRANSCRIPTION AT REDUCED ENZYME CONCENTRATIONS

In order to assess the productivity of T7 RNA polymerase at reduced enzyme concentrations, a series of transcription reactions were performed using solution phase DNA templates. The time dependent increase of total RNA product formation was measured in duplicate at three separate enzyme concentrations [23 μg/ml, 11 μg/ml and 5 μg/ml] over the course of four hours. The quantitated results are expressed in graphical form in figure 3 (data represent average values from two separate experiments). This graph measures the cumulative total [α-^{32}P]-UTP incorporation as a function of time. Each curve is characterized by an initial linear phase which gradually plateaus yielding a classic Monod curve. The slope of the initial enzyme velocities decreases proportionately (to a first approximation) with decreasing enzyme concentration. Interestingly, at longer timepoints, the overall productivity does not decrease proportionately with enzyme concentration. This suggests that the enzyme remains active for extended reaction times, ultimately yielding increased RNA production per polymerase molecule. This is evidenced by the fact that at 23 μg/ml the extent of UTP conversion reaches fifty percent, whereas at 6 μg/ml thirty percent conversion is attained. Since the only variable in this study is enzyme concentration, other factors can be ruled out. Consequently, these results suggest that T7 RNA polymerase productivity can be increased by reducing the enzyme concentration and allowing the reaction to proceed for extended time periods. This is significant because standard transcription protocols recommend enzyme concentrations ranging from 0.5 to 0.1 mg/ml.

PREPARATIVE, SEMI-BATCH TRANSCRIPTION REACTIONS

In order to further demonstrate the application of immobilized DNA templates, we conducted preparative scale (1.0 ml) transcription reactions in a stirred-cell reactor operated in semi-batch mode (as described in the methods section). Unfortunately, the results obtained from these semi-batch transcription experiments are merely qualitative. In general, RNA products were detected only in the first two timepoints (one and two hours, respectively) when immobilized templates were used with one initial enzyme dose (data not shown). This finding is in sharp contrast to reactions performed with solution phase templates, where RNA products were present through the fourth hour of the reaction (data not shown). Solution phase reactions appeared to yield more product than immobilized templates, challenging the results obtained form small-scale quantitative comparisons. Nevertheless, transcriptional activity from immobilized DNA templates could be rescued by adding a second dose of fresh enzyme midway through the reaction sequence (data not shown). In addition, immobilized templates were active in successive rounds of semi-batch transcription reactions, despite somewhat lower apparent yields. This result is significant because it further verifies the stability and activity of immobilized DNA templates in semi-batch transcription applications.

Immobilized DNA templates are retained in the stir-cell reactor, whereas solution phase templates (molecular weight 30 kilodaltons) are small enough to pass through the 100 kilodalton ultrafiltration membrane into the product stream. Consequently, we expected that the solution phase templates would wash-out of the reactor during the course of the semi-batch transcription sequence. Since RNA product was detected up through four hours with a single dose of enzyme and up to eight hours with a second enzyme supplement, it does not appear that the solution phase templates wash-out of the reactor. This may be due in part to the specific binding of RNA polymerase to template molecules during transcription, forming complexes which are selectively retained by the membrane. Alternatively, solution phase templates may in fact wash-out during the course of the reaction. In this case, products observed at later timepoints merely reflect RNA which was transcribed early on but not removed until subsequent timepoints. More experiments are required to distinguish between these two separate possibilities.

REFERENCES

[1]Saenger, W. (1984) In *Principles of Nucleic Acid Structure* (C. R. Cantor, ed.) Chapters 6, 8 and10. Spinger-Verlag, New York, NY.

[2]Cech, T. R. (1987) The chemistry of Self-splicing RNA and RNA enzymes, *Science* **236**:1532-1539.

[3]Chamberlin, M. and Ryan, T. (1982) Bacteriphage DNA-dependent RNA polymerases, In *The Enzymes* (P. D. Boyer, ed.), 3rd ed., Vol. 15 pp. 87-108, Academic Press, New York, NY.

[4]ibid.

[5]Davanloo, P., Rosenberg, A. H., Dunn, J. J. and Studier, F. W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase, *Proc. Nat. Acd. Sci.. USA* **81**:2035-2039.

[6]Milligan, J. F., Groebe, D. R., Witherell, G. W. and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nuc. Acids Res.* **15**: 8783-8798.

[7]ibid.

[8]Chamberlin, M. and Ryan, T. (1982) Bacteriophage DNA-dependent RNA polymerases, In *The Enzymes* (P. D. Boyer, ed.), 3rd ed., Vol. 15 pp. 87-108, Academic Press, New York, NY.

FIGURE 1

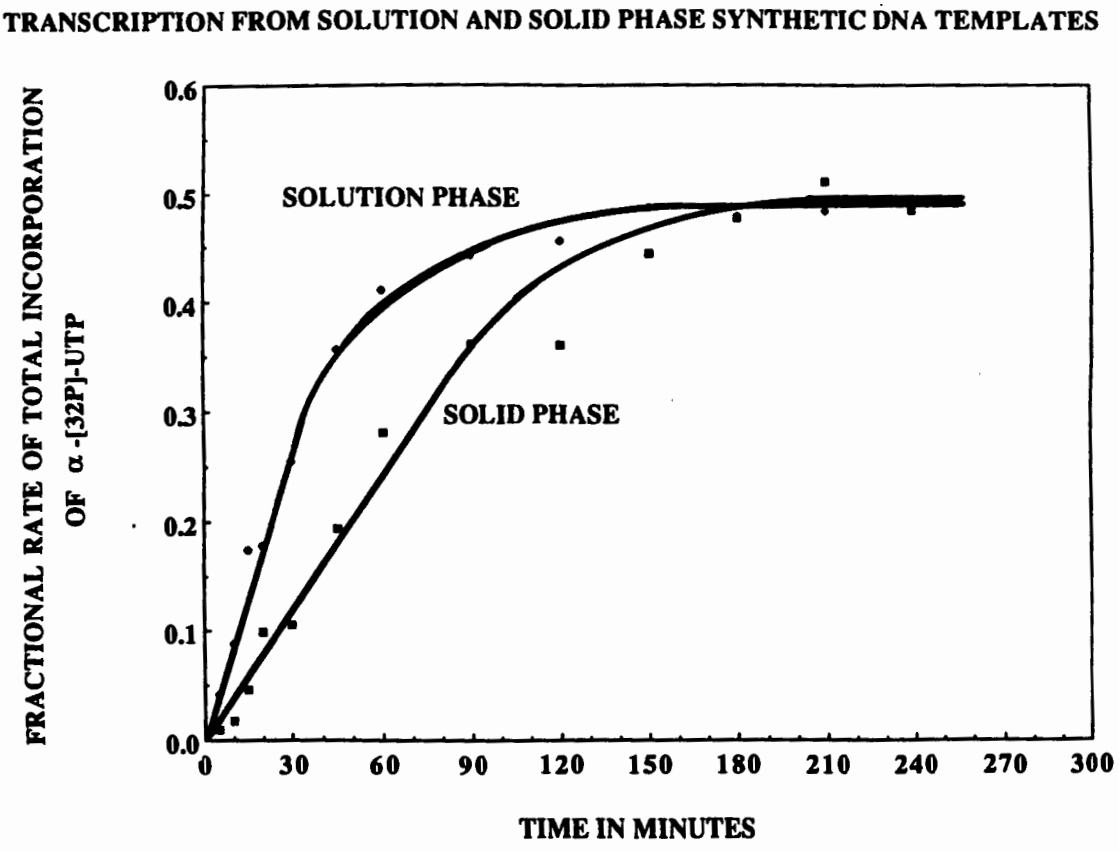


FIGURE 2

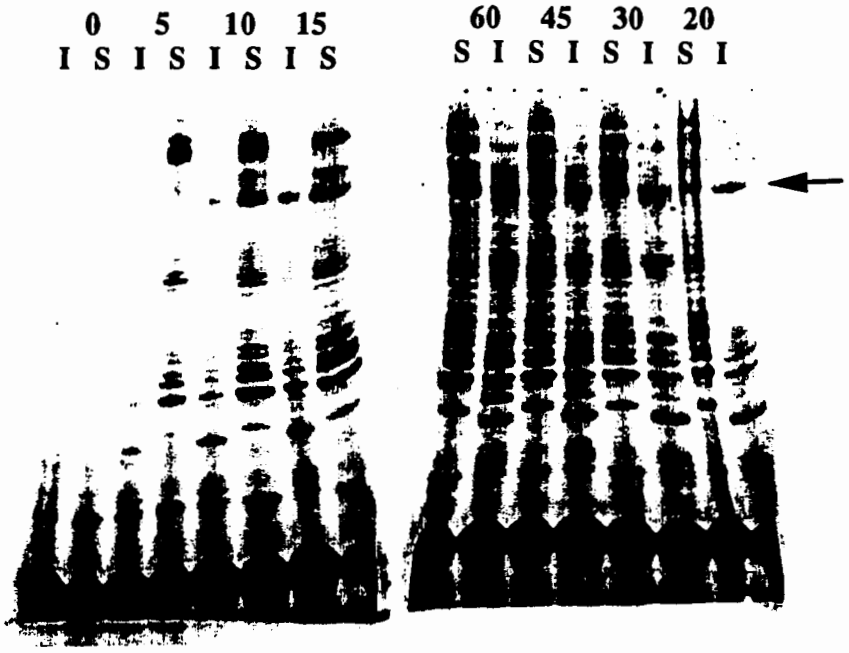
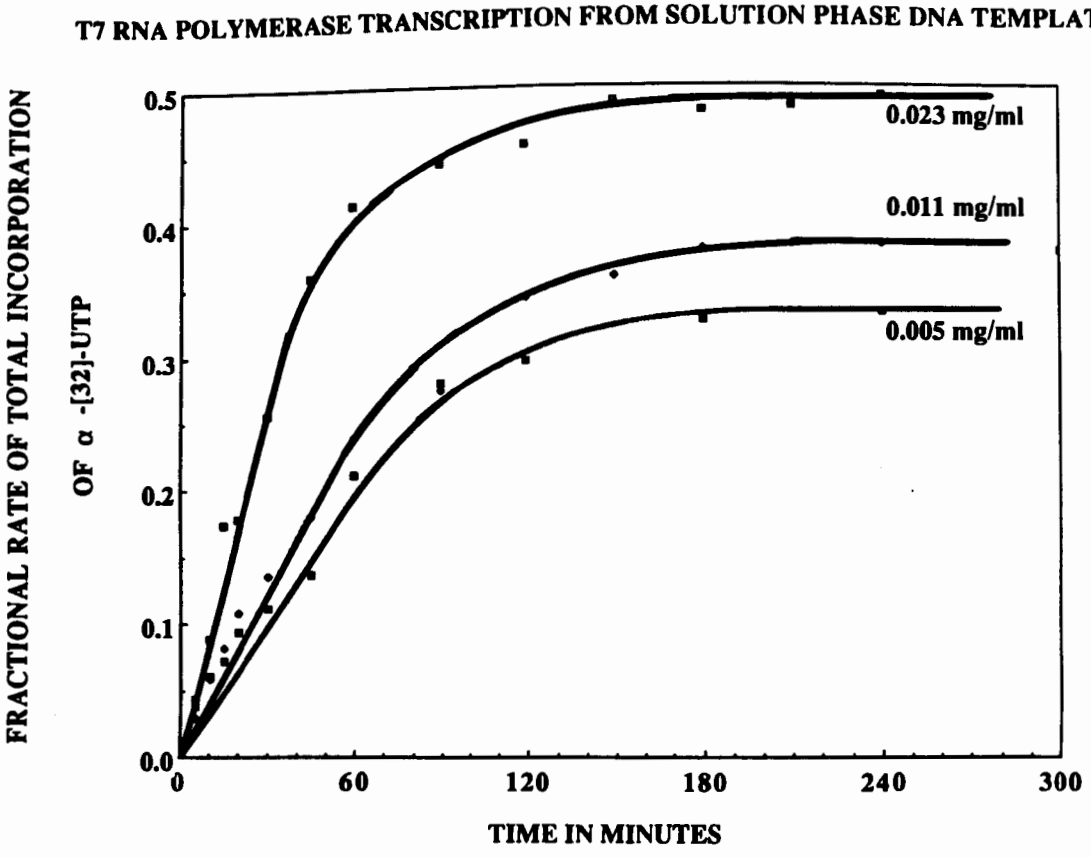


FIGURE 3



DEVELOPMENT OF A CELSS BIOREACTOR: OXYGEN TRANSFER AND MICROMIXING IN PARABOLIC FLIGHT

P.E. Villeneuve[1], K.S. Wenger[1], B.G. Thompson[2], T. Kedar[3] and E.H. Dunlop[1]

[1]Colorado State University, Dept. of Ag & Chem. Engineering, Fort Collins, CO 80523
[2]Alberta Research Council, Biotechnology Dept., Edmonton, Alberta, Canada T6H 5X2
[3]Strathclyde University, Dept. of Pure and Applied Chemistry, Glasgow, Scotland

ABSTRACT

The gas exchange portion of a phase-separated loop bioreactor was tested with respect to oxygen mass transfer and micromixing in accelerations of 0.01g, 1g, and 2g. A plot of the overall mass transfer coefficient versus gravity indicates the rate of oxygen transfer does not change as a function of acceleration. Also, it was determined that the micromixing did not exhibit significant changes in the various gravitational fields. These observations indicate the loop bioreactor should function independent of acceleration.

INTRODUCTION

Future human development of space will require processes that involve mixing and mass transfer considerations similar to those used on Earth. One such process involves the cultivation of microorganisms. Several designs of bioreactors for space applications have been proposed \1-4\. Villeneuve and Dunlop have developed a phase-separated gravity-independent loop bioreactor in order to carry out fermentations in space \4\. The bioreactor will support both carbon recycling algae and heterotrophic microorganisms, both of which exhibit normal growth in space \5\. The bioreactor is viewed as a back-up system or adjunct to a on-board plant and animal carbon cycling schemes \6-8\.

We have demonstrated that this system prevents phase mixing and transfers oxygen with high efficiencies at unit gravity \4\. However, it has not been tested in low gravitational environments. It is not known if this reactor will maintain the same rate of oxygen transfer and the same micromixing conditions in space. Both micromixing and the rate of oxygen deliver can significantly influence microbial growth. It has been demonstrated the micromixing can potentially alter carbon conversion efficiencies \9\. The rate of oxygen transfer directly influences the reactor size \4\. Therefore, both the overall mass transfer coefficient and the micromixing was examined in hypergravity and microgravity on the NASA KC-135.

METHODS AND MATERIALS

Generation of microgravity. Experiments were carried out on the NASA KC-135. This is a modified four-engine in-flight fueling transport jet. Each flight sequence consisted of forty parabolas. Each parabola lasted about 75 seconds, of which about 15-20 seconds are at 0.01g or less, followed by a 1.8-2.0g period. Reduced acceleration is generated at the top of each parabola when the acceleration vector, generated by the plane, is of the same magnitude as the acceleration of unit gravity but in the opposite direction.

Micromixing test reaction. The chemical test reaction used to make the micromixing measurements was the reaction of 1-naphthol with diazotized sulfanilic acid \10-12\. The use of this reaction as a micromixing test reaction is well-documented in these references. The reaction can be adequately described by the competitive, consecutive scheme in equation (1) \13\.

$$A + B \rightarrow R + B \rightarrow S \qquad (1)$$

The product distribution was analyzed spectrophotometrically using the method described by Wenger *et al.* \13\. One consideration that had to be made in the design of the reactor was to ensure that the reaction was complete

during the residence time of the reactor. Using the second order rate constants measured by Bourne *et al.* \11\, the reactor was modeled as a series of CSTR's. The model indicates the reaction is 99.0% complete 0.1 m down the reactor.

Micromixing experimental apparatus. A schematic of the experimental reactor is given in Figure 1. The reactor was constructed of glass with an inner diameter of 0.01m. Two turbulence triggers, 3 mm diameter glass rods, were placed upstream of the injection point to ensure equilibrium turbulence.

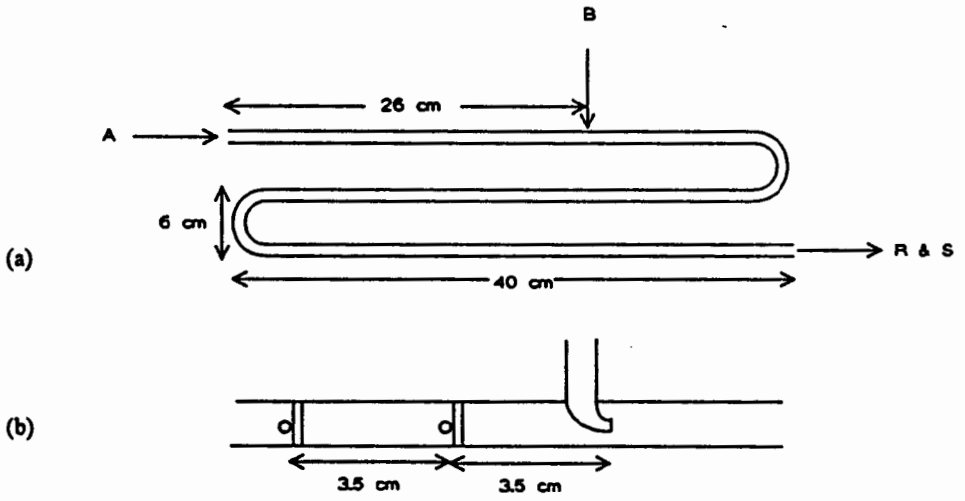


Fig. 1. (a) Schematic diagram of the KC-135 micromixing test apparatus experimental reactor. (b) Expanded view of the turbulence triggers.

The buffered 1-naphthol solution (A) initially flowing in the pipe encountered the B solution which was injected isokinetically through a 0.0032 m I.D. injection pipe. The samples were drawn by the experimental operator into Vacutainers. Reagent solutions A and B were initially stripped of oxygen using nitrogen and were stored in collapsible polyethylene bags during the flight. Waste product solution was similarly stored in collapsible bags. The reaction was carried out in continuous mode with $F_A/F_B = 10$. Two flight experiments were conducted, at Reynolds numbers of 5,300 (Flight 1) and 8,600 (Flight 2).

Oxygen transfer experimental apparatus. Deoxygenated distilled water was pumped into the loop reactor on a continuous basis at a constant 60 ml/min with a peristaltic pump (Figure 2). The loop reactor consisted of

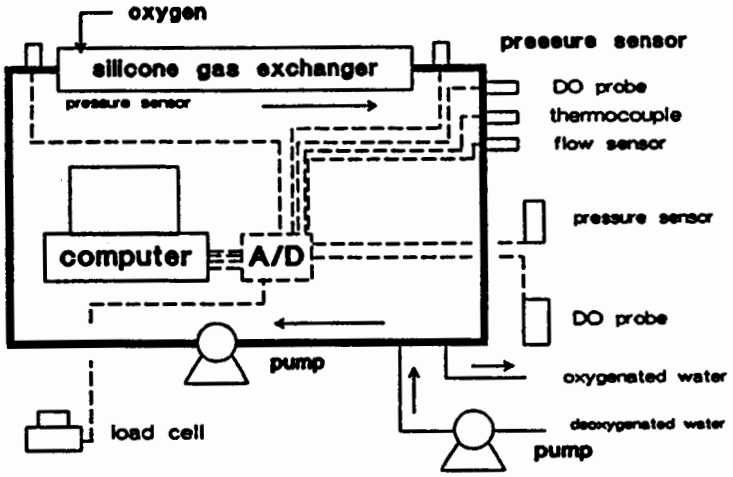


Fig. 2. Schematic of the KC-135 oxygen transfer test apparatus.

stainless steel pipe, fittings and silicone tubing. The silicone oxygen exchanger provided surface area for oxygen exchange. Twenty liter high density polyethylene collapsible bags provided supply and return reservoirs for water. Fifteen liters of distilled water in the supply tank was deoxygenated with nitrogen gas two hours prior to each flight. Analog signals were converted to digital and logged to a floppy disk every two seconds throughout the

entire flight. Monitored signals included dissolved oxygen, temperature, pressure drop across the reactor, weight (load cell) and cabin air pressure. The re-circulation flow-rate was held constant at Re = 4,800 for the first twenty parabolas of each flight. The re-circulation flow-rate was stepped up to generate Reynolds numbers of 6,600 and 8,300 during the last twenty parabolas of each flight. Five parabolas were allowed to pass before the flow-rate was changed. YSI 5739 polarographic dissolved oxygen (DO) probe with 23 thick FEP Teflon membrane was used as it was reliable, accurate, and it has a relatively quick response time.

Post flight analysis. Subsequent analysis involved calculating the overall mass transfer coefficient and a Sherwood number every 2 seconds throughout the flight. Multiplying DO by feed-rate and dividing by reactor volume provided a oxygen transfer rate. This value was used to calculate an apparent K_l from the following equation: $k_l a = \frac{rate}{(c^* - c)}$. Saturated dissolved oxygen concentration is obtained from Henry's law and the bulk concentration was measured with the DO probe. The Sherwood number was calculated as: $Sh = \frac{k_l d}{D_{O_2}}$.

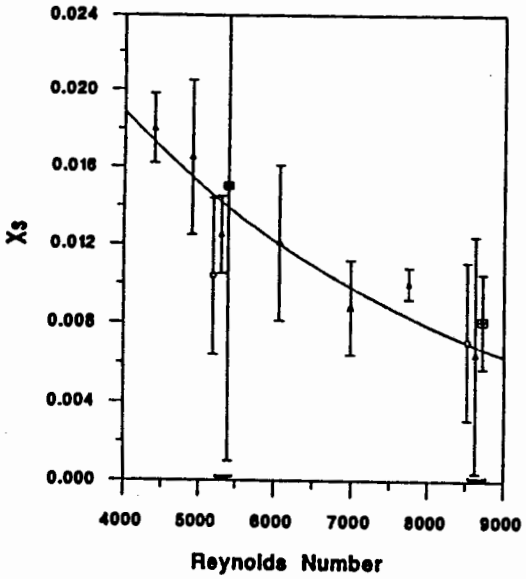


Fig. 3. X_s as a function of Reynolds number. Triangles ground based, diamonds low g and squares high g.

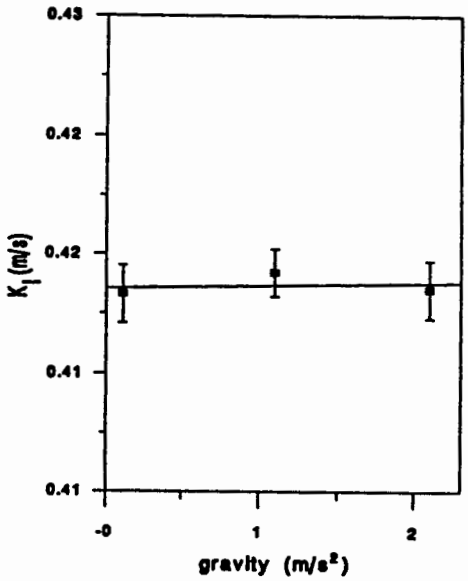


Fig. 4. Mass transfer coefficient versus gravity

RESULTS AND DISCUSSION

Micromixing results. Figure 3 shows X_s as a function of Reynolds number for the high and low g portions of the parabola compared with ground based data. Ground based data are based on 10 different samples taken at the same conditions as the flights. Within the reproducibility of the analysis, it is apparent that gravity had no significant effect on the product distribution. Analysis of variance between the low g, ground, and high g data for the two flights indicated that the differences were not significant at the 5% level. Analysis of variance for the ground data as a function of Reynolds number was significant at the 1% level. As the large amount of scatter for the high g portion of Flight 1 was due to two outlying points, it is possible that more data would have revealed a significant effect.

Oxygen transfer results. The data of the constant flow-rate experiment are presented in Figure 4. Fluctuations are within the measured standard deviation. A total lack of trends and periodicity in the K_l would indicate the oxygen mass transfer did not vary with gravitational force at a constant flow-rate. Plots of Sherwood versus Reynolds number were generated for the variable flow rate portion of each flight. Slopes of these plots along with slopes generated from ground based Sherwood versus Reynolds plots are presented in Table 1.

TABLE 1 Sherwood versus Reynolds slopes

replicates	ground studies	high gravity	low gravity
r_1	0.056	0.112	0.126
r_2	0.079	0.019	0.019
r_3	0.035	0.043	0.045
mean	0.056	0.058	0.063
STD	0.0022	0.0048	0.0055

Changing the re-circulation rate during the last 20 parabolas provided the opportunity to investigate if an "gravity effect" exists at higher flow rates. This

part of the experiment was also designed to test the reactor under potential mission conditions. Slopes of Sherwood versus Reynolds number generated in-flight and on the ground were examined and compared statistically. The statistical analysis pointed out the three treatments are equal at a 95% confidence interval.

CONCLUSIONS

Under the specified conditions, it appears that gravity does not affect oxygen mass transfer rates or micromixing. Therefore, the bioreactor should function in the same manner with respect to oxygen delivery and micromixing in space as it does on earth.

REFERENCES

1. P.K. Seshan, G.R. Petersen, B. Beard, C. Boshe and E.H. Dunlop, Design Concepts for Bioreactors in Space, *NASA TM* 88215 (1986).

2. J.V. Mayeux, Influence of Zero-G on Single-Cell Systems and Zero-G Fermenter Design Concepts, *NASA TM* X-58191 (1977).

3. G.R. Petersen, P.K. Seshan and E.H. Dunlop, Phase Separated Membrane Bioreactor: Results From Model System Studies, *Adv. Space Res.* 9, #8, 185 (1989).

4. P.E. Villeneuve and E.H. Dunlop, Evolution of a Phase Separated Gravity Independent Bioreactor, *Adv. Space Res.* 12, #5, 237 (1992).

5. F.K. Gmunder and A. Cogoli, Cultivation of Single Cells in Space, *Applied Microgravity and Technology* 3, 115(1988).

6. M.L. Shulere, D. Nafis and E. Sze, The Potential Role of Aerobic Biological Waste Treatment in Regenerative Life Support Systems, *ASME 81-ENAs*-20 (1981).

7. D.W. Jenkins, Bioregenerative Systems, *NASA SP*-165, 1 (1968).

8. R.W. Krauss, Closed Ecology in Space From a Bioengineering Perspective, *Life Sciences and Space Research* 17, 13 (1979).

9. E.H. Dunlop and S.J. Ye, Micromixing in Fermentors: Metabolic Changes in *Saccharomyces cerevisiae* and Their Relationship to Fluid Turbulence, *Biotechnology and Bioengineering* 36, 854 (1990).

10. J.R. Bourne, F. Kozicki and P. Rys, Mixing and Fast Chemical Reaction - I, *Chem. Eng. Sci.* 36, #10, 1643 (1981).

11. J.R. Bourne, C. Hilber and G. Tovstiga, Kinetics of the Azo Coupling Reactions Between 1-Naphthol and Diazotized Sulfanilic Acid, *Chem. Eng. Commun.*, 37, 293 (1985).

12. J.R. Bourne, O.M. Kut, J. Lenzner and H. Maire, Kinetics of the Diazo Coupling Between 1-Naphthol and Diazotized Sulfanilic Acid, *Ind. Eng. Chem. Res.* 30, 1385 (1990).

13. K.S. Wenger, I. Macgilp, and E.H. Dunlop, An Investigation of the Chemistry of a Diazo Micromixing Test Reaction, *AIChE J.* 38, #7, 1105 (1992).

NOMENCLATURE

A	1-naphthol or its reagent solution	K_l	overall mass transfer coefficient (m/s)
B	Diazotized sulfanilic acid	a	gas liquid interface area per liquid volume (m^{-1})
F_A F_B	flow rate of A & B (m^3/sec)	C*	saturated dissolved oxygen concentration (mg/l)
R	4-(1,4-sulfophenylazo) 1-naphthol	C	bulk dissolved oxygen concentration (mg/l)
S	2,4 bis (1,4-sulfophenylazo) 1-naphthol	Sh	Sherwood number
X_s	product distribution	d	characteristic length (m)
D_{O2}	oxygen diffusivity in water (m^2/s)	g	gravity (m/s^2)

Scale-up of Dexter murine bone marrow cultures utilizing a three-dimensional fiberglass support matrix

John G. Highfill, Paul Todd, Steve Haley, and Dhinakar Kompala
Department of Chemical Engineering, University of Colorado-Boulder, Boulder, CO 80309-0424
phone: 303-492-6350; fax: 303-492-4341

Introduction

The *ex vivo* expansion of bone marrow cells has important uses in bone marrow transplantation therapies crucial for various cancer treatments including myelogenous leukemia[6] and poor prognosis breast cancer[7]. The current method of treatment for these diseases is typically high-dose chemotherapy, followed by transplantation of either cryopreserved autologous marrow or allogeneic marrow obtained from a compatible donor. These treatments typically require 14 days or longer for restoration of the immune system. Thus the development of bone marrow expansion technology is crucial to reduce this time in which the immune system is extremely susceptible to infection. Future applications of marrow expansion may also include the purging of acute myeloid leukemia cells existing within the bone marrow[1], and the use of gene therapy on early stem cells to alter genetic defects[3].

Bone marrow cells are commonly cultured on the flat two-dimensional surface of tissue culture dishes. Dexter and coworkers developed a long-term murine bone marrow culture system (Dexter cultures) which gives rise to myeloid cells[2]. In the first 7-10 days following inoculation, an adherent layer of cells grows to confluency. This cell population is commonly called *stromal* cells, and consists of numerous cell types including endothelial cells, adipocytes, macrophages, and fibroblasts[4]. These cells secrete cytokines and growth factors which stimulate differentiation of hematopoietic stem cells into more blood mature cells. However, the hematopoietic cells do not typically reach complete maturity and the more immature cells, sometimes called progenitor cells, form hematopoietic colonies within 14-21 days. As the colonies grow in size, some cells detach from the colony into the culture medium suspension and can be harvested for analysis. Our research on bone marrow culture scale-up

utilizes Dexter mouse bone marrow culture systems as models to develop further improvements in human bone marrow cultures.

The objective of this work is to demonstrate scale-up of murine bone marrow culture systems from monolayer tissue culture dishes (t-25) to a petri dish containing two fiberglass disks, giving a 22-fold increase in surface area available for cell growth. This system provides a more natural three-dimensional environment for cell-cell interactions.

Cell growth cannot be monitored efficiently in attachment-dependent culture systems. Therefore, glucose consumption rates were analyzed as indications of cellular metabolic activity. Suspended cells were counted and analyzed morphologically to determine cell type. Finally, attached cells were enumerated using nucleus counts.

Materials and Methods

Reagents and Culture Media

Minimum essential medium alpha (α-MEM), penicillin/streptomycin, phosphate buffered saline solution (PBS), and horse serum were purchased from Gibco (Gaithersburg, MD); 2% gelatin, fibronectin, pluronic F-68, and hydrocortisone were purchased from Sigma (St. Louis, MO).

Dexter culture medium consisted of α-MEM containing 50 units/mL penicillin, 50 μg/mL streptomycin, 10^{-6}M hydrocortisone, and 20% horse serum. For some experiments, the culture medium was supplemented with pluronic F-68 up to 0.1% (w/v).

Establishment and Monitoring of Monolayer Cultures

Dexter monolayer cultures were initiated by harvesting bone marrow from the femurs and tibiae of 9-10 week old Balb/c mice obtained from Harlan Sprague Dawley (Indianapolis, IN). The cells harvested from one mouse (~5 x 10^7 nucleated cells) were divided into four 25-cm^2 tissue culture flasks (Becton Dickinson, Lincoln Park, NJ) in 8 ml Dexter culture medium per flask. The cultures were typically fed by removing 4-5 mL old medium and replacing with the same volume of fresh medium after 1 week, and subsequently every 3 or 4 days afterward. Dexter cultures were incubated at 33°C and 6% CO_2 in air.

Harvested cells were counted on a hemacytometer and viability was assessed based on trypan blue (Sigma) exclusion. The harvested medium was also used to determine glucose and lactate concentrations with a Model 2000 Glucose/l-Lactate analyzer (Yellow Springs Instruments, Yellow Springs, OH).

Cell typing was performed by plating the harvested cells onto a microscope slide using a cytocentrifuge (Cytospin I, Shandon, Inc., Pittsburgh, PA) at 700 rpm for 10 minutes. The slide was then submerged in a solution of Wright-Giemsa stain (Fisher Scientific, Pittsburgh, PA) for 15 seconds, placed in deionized water for 15 seconds and dipped in another deionized water solution 10 times. After air drying, the stained cells were scored as myeloid, monocyte/macrophage, blast, or *other*.

Three-dimensional culture on fiberglass disks

Fiberglass mat (OCMat 3200) with a mean fiber diameter of 13.5μm was obtained from Owens-Corning Fiberglas (Toledo, OH) and cut into 58 mm disks. A 19 mm annular hole was cut out of the middle each disk. The disks have a thickness of 0.7 mm and 280 cm^2 theoretical surface area available for growth. Two of these disks were stacked in an untreated petri dish (Fisher Scientific) and inoculated with freshly harvested bone marrow cells. Feeding schedules for these cultures were based on the glucose concentration of the harvested medium, which was not allowed to fall below 0.10 mg/mL.

Experiments were carried out to determine the effects of the presence of attachment factors on cell growth. The fiberglass disks were coated with fibronectin (10 μg), gelatin (10 μg), and a combination of fibronectin and gelatin (10 μg each) in untreated petri dishes. Each attachment factor was diluted in PBS. After coating, the disks were allowed to air dry overnight, and were washed two times with PBS and transferred to new untreated petri dishes for inoculation.

Analysis of attached cells

Fiberglass disks taken from petri dish cultures and the bioreactor were analyzed for attached cell content by performing nucleus counts[5]. The disks were submerged in a solution containing 0.1 M citric acid, 1% triton X-100, and 0.1 mg/mL crystal violet. Next, they were agitated on a platform stirrer for 15 minutes and the stained nuclei counted on a hemacytometer.

Results and Discussion

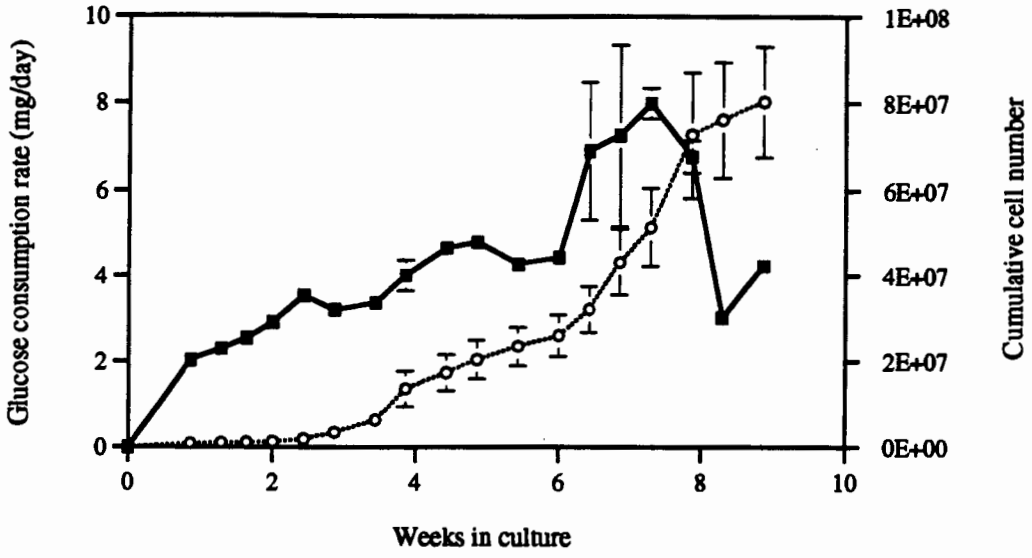
Glucose consumption rates in monolayer cultures

An indication of cell growth kinetics is given by the glucose consumption rates, which were computed by differentiating cumulative glucose consumption data and are shown in Figure 1 (black squares). The consumption rate increased dramatically for the first 7-10 days, during which time the stromal cells are growing to confluency. Between 10 and 21 days, stromal cell growth declines and hematopoiesis begins, and although the consumption rate still increases, the rate of increase is slower. During this phase of the culture, the glucose consumption rate includes the dormant stromal cells, the attached progenitor cells and the more differentiated cells in suspension. During the sixth week, feed volumes were increased slightly to 6 mL. This initially resulted in an increase in the glucose consumption rate. However, suspended cells began to wash out, leading to a subsequent decrease in glucose consumption rate. The rate of glucose consumption reached a maximum of 8.0 mg/day.

Cell harvest and morphological staining

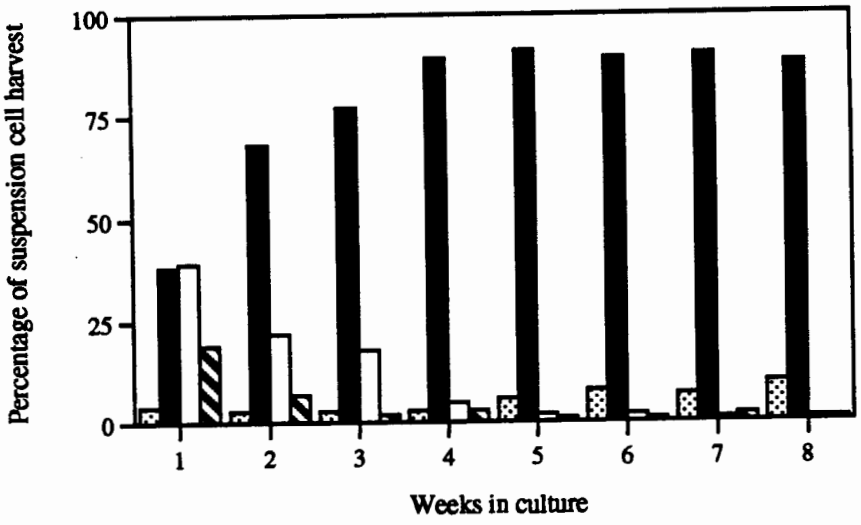
The cumulative cell number harvested from suspension, also shown in Figure 1 (open circles) further illustrates these effects. Very few cells were harvested until week 3, when hematopoiesis had begun. From week 4 through 6, the typical cell concentration was 10^6 cells/mL. After the feeding volume was increased at week 6, the harvested cell concentration increased up to 4-fold. At the end of the culture period, the expansion (ratio of total cells harvested to cells inoculated) for these cultures was 6.4. The error bars indicate the standard deviation of three cultures.

Figure 1: Glucose consumption rates and cumulative cell harvest in monolayer cultures



Differential analysis of cells harvested from suspension and stained with Wright-Giemsa stain is shown in Figure 2 (black = granulocytes, white = other, dotted = monocyte/macrophages, and striped = blasts). These results show that the fraction of granulocytic cells increased until week 5 and then remained above 85% of the harvested population. Also at week 5, the proportion of macrophages began increasing, reaching 10% of the cell population by week 8. The large fraction of *other* cells at week 1 and 2 was primarily lymphoid and mature erythrocytes which disappeared by week 3.

Figure 2: Differential staining of cells harvested from monolayer cultures

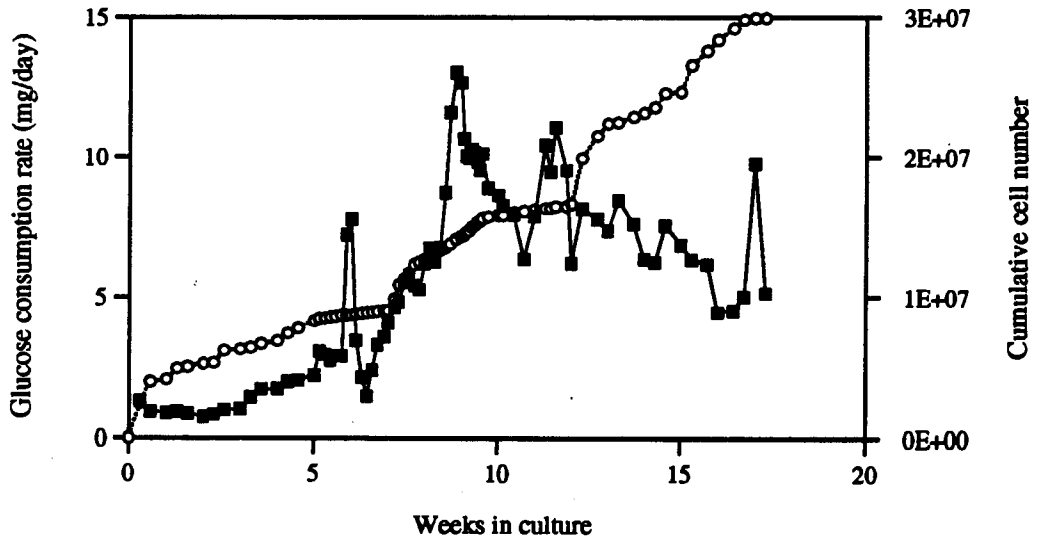


Fiberglass disk cultures

A set of three cultures was grown on 2 fiberglass disks placed in an untreated petri dish. The inoculum of these cultures consisted of 5×10^7 freshly harvested bone marrow cells. They were also reseeded with 2.75×10^7 cells during the sixth week. The glucose consumption rates for one of these cultures are shown in Figure 3 (black squares). The culture presented in this graph is the untreated control, which had the highest maximum glucose consumption rate. Following the reseeding of this culture with fresh bone marrow , the glucose consumption rate increased dramatically, then decreased even more dramatically during the sixth and seventh weeks, before climbing steadily to a maximum during week 8 and declining steadily thereafter. The maximum glucose consumption rate for this culture was 12.3 mg/day.

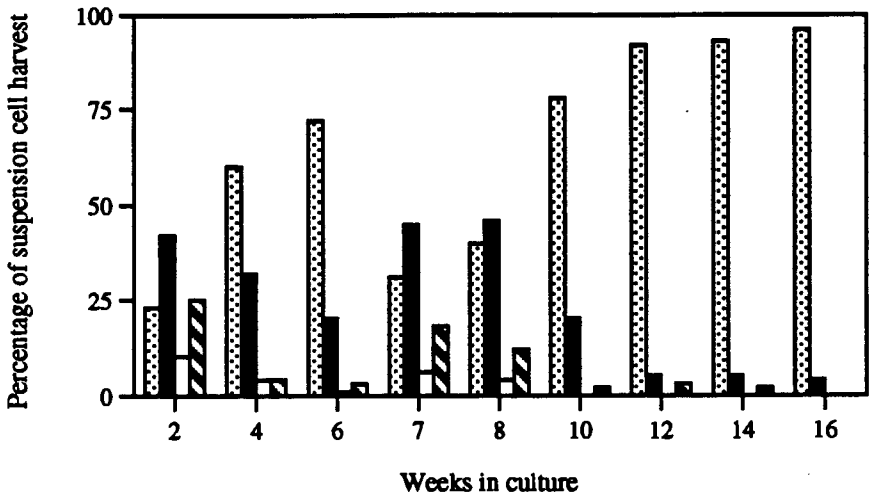
The cumulative cell numbers harvested from suspension for this culture are also given in Figure 3 (open circles). During week 7, when larger volumes of medium were fed in order to keep the glucose concentration above 0.10 mg/mL, the cumulative cell number harvested increased significantly. Another increase occurred during week 12 when the disks were washed thoroughly by vigorous pipetting of the culture medium. Suspended cell concentrations in the harvested medium were less than 10^5 cells/mL, indicating that active hematopoiesis did not occur.

Figure 3: Glucose consumption rates and cumulative cell harvest in fiberglass disk cultures



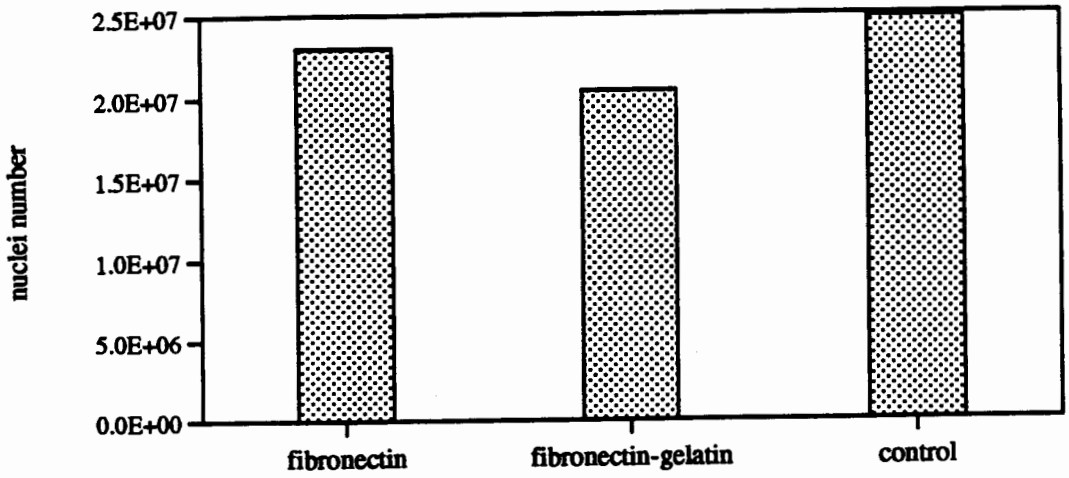
Differential analysis of the cells harvested from suspension and stained with Wright-Giemsa stain did not result in the expected cell types. Figure 4 shows the scoring of these cells (black = granulocytes, white = other, dotted = monocyte/macrophages, and striped = blasts). At week 2, the granulocytes were still less than 20% of the entire population, whereas the macrophage population was greater than 20%, much higher than that observed for monolayer cultures. From week 2 to 6, the macrophages increased to 75% of the total population. Following reseeding with fresh marrow, the macrophages again became the predominant cell type in suspension for the remainder of the culture period. In contrast, granulocytic cells decreased from week 2 to 6, and though replenished from reseeding, they decreased again to less than 5% from week 8 to the end of the culture.

Figure 4: Differential staining of cells harvested from fiberglass disk cultures



Following termination of the cultures, the number of cells attached to the fibers was analyzed using crystal violet nucleus counts. The results presented in Figure 5 surprisingly indicate that the untreated control produced the largest number of attached cells, although the cultures with adsorbed attachment factors produced nearly as many cells. This effect is considered to be a result of the high concentration of serum present in the medium (20% horse serum) which provides proteins necessary for cell attachment, eliminating the need for exogenously added attachment factors.

Figure 5: Nucleus counts for attached cell content in fiberglass disk cultures



A summary of the scale up achieved is presented in Table I. Maximum glucose consumption rates for all cultures were normalized according to surface area available for cell attachment and growth. The results indicate that a specific glucose consumption rate of more than 300 $\mu g/day/cm^2$ is attainable in Dexter monolayer cultures. Thus, the fiberglass disk cultures have the potential of utilizing 168 mg/day, or more than 10 times the maximum glucose consumption rate actually achieved. These consumption rates were not achieved in fiberglass disk cultures for two reasons. First, the stromal cells did not reach confluency. This was verified visually by observation of the fiberlgass disks under a phase contrast microscope. The cells appeared to prefer to grow in areas of crosslinked fibers, although growth along a single fiber was common. In fact, the calculated surface area available for cell growth is lower than observed because the cells often grew in three-dimensional clusters in which very little or no attachment was made with the fibers. This is also expected during colony formation. The second reason that predicted glucose consumption rates were not achieved was that hematopoiesis did not occur to any observable extent. High concentrations of cells ($> 10^5$ cell/mL) in suspension were not produced, and morphological analysis of cells harvested from suspension

revealed that the majority of these were macrophages, which are also known to comprise a portion of the stromal cell layer[4]. Interestingly, Sardonini and Wu[8] reported that monocyte/macrophages were the majority of the population after only 14 days in their scale up attempts of human bone marrow cells in spinner flask suspension cultures, and increased to 89% by 28 days in culture. This was attributed to the presence of IL-3, granulocyte-macrophage colony stimulating factor (GM-CSF), and stem cell factor (SCF), which was exogenously added to the culture medium. However, our cultures were not supplemented with any cytokines. Therefore, the production of monocyte/macrophages may be a systemic result.

Table 1: Summary of scale-up

	maximum glucose consumption rate (mg/day)	specific glucose consumption rate ($\mu g/day/cm^2$)	total cell number harvested	ratio of viable cells harvested to cells inoculated
monolayer culture (t-25)	7.98	319.2	8.0×10^7	6.4
fiberglass disk culture	12.3	22.0	5.5×10^7	0.71
in suspension			3.0×10^7	0.39
attached			2.5×10^7	0.32

References

1. Chang, J., Allen, T. D., Dexter, T. M. 1989. Long-term bone marrow cultures: Their use in autologous marrow transplantation. Cancer Cells 1:17-23.
2. Dexter, T. M., Allen, T. D., Lajtha, L. G. 1977. Conditions controlling the proliferation of hemopoietic stem cells *in vitro*. J. Cell. Physiol. **91**: 335-344.
3. Gutierrez, A. A., Lemoine, N. R., Sikora, K. 1992. Gene therapy for cancer. The Lancet **339**: 715-721.
4. Johnson, A., Dorshkind, K. 1986. Stromal cells in myeloid and lymphoid long-term bone marrow cultures can support multiple hemopoietic lineages and modulate their production of hemopoietic growth factors. Blood **68**: 1348-1354.
5. Lin, A. A., Nguyen, T., Miller, W. M. 1991. A rapid method for counting cell nuclei using a particle sizer/counter. Biotech. Tech. **5**: 153-156.
6. McMillan, A. K., Goldstone, A. H., Linch, D. C., Gribben, J. G., Patterson, K. G., Richards, J. D. M., Franklin, I., Boughton, B. J., Milligan, D. W., Leyland, M., Hutchinson, R. M., Newland, A. C. 1990. High-dose chemotherapy and autologous bone marrow transplantation in acute myeloid leukemia. Blood **76**: 480-488.
7. Peters, W. P., Shpall, E. J., Jones, R. B., Olsen, G. A., Bast, R. C., Gockerman, J. P., Moore, J. O. 1988. High-dose combination alkylating agents with bone marrow support as initial treatment for metastatic breast concer. J. Clin. Oncol. **6**: 1369-1376.
8. Sardonini, C. A., Wu, Y.-J. 1993. Expansion and differentiation of human hematopoietic cells from static cultures through small-scale bioreactors. Biotech. Prog. **9**: 131-137.

MODELING AND ESTIMATION OF STATES OF RECOMBINANT FERMENTATIONS USING NONLINEAR INPUT/OUTPUT MODELS

Victor M. Saucedo and M. Nazmul Karim
Department of Ag. and Chemical Engineering
Colorado State University
Fort Collins, Colorado

ABSTRACT

This research investigates a reliable technique for state estimation based on observable information using models that allow us to more easily and accurately identify the parameters of the model. The basic approach for modeling relies on the use of nonlinear input/output models. Volterra series, a nonlinear input/output model, is used for modeling practical cases using an update of kernels through a moving window or the addition of an autoregressive term. Two important concepts are introduced for model selection. One characterizes the quality of the input (the periodogram), and the other selects the best model for a number of parameters (the AIC). In order to obtain the best estimates when using these input/output models, the Kalman filter recursions were used coupled with the state representation of the model.

In this research, models are used to identify recombinant Escherichia coli fermentations. Batch and cell recycle experiments were performed for ethanol production. Here, it is shown that when Volterra series includes an autoregressive term, it can identify the dynamic behavior of the process reasonably well.

1 Introduction

Models are required for a better understanding of the system and for engineering purposes such as design, optimization and control. The typical modeling approach is based on differential equations coming from different laws such as energy or mass balances. Nevertheless, this area faces a big roadblock because of ambiguity when it comes to determine the proper values from experimental results. Another important aspect is that the parameters in the selected model sometimes change significantly during the process. This is known as a time variant system. These problems are more easily seen when estimating parameters of biological processes. Model parameters may change over long periods of time due to physiological changes in the microorganisms.

In order to optimize continuous or semi-continuous processes one has to deal with a lot of problems: first, measure the desired variables; second, determine the behavior of the system; and third, determine the right input to the system for the control law. The measurement of variables important for optimization purposes is many times impossible to obtain on-line. Efforts are being made in the development of reliable physical sensors and algorithms to estimate and predict the measurement. However, no matter how close scientists get to the development of appropriate sensors, it will always be worthwhile to continue the development of theoretical estimators and filters, because these will not only improve the available measurements (filter from noise), but offer additional capabilities such as prediction, which is useful for on-line process control.

One approach for modeling nonlinear time variant systems is the use of Volterra series, an input/output approach. The selection of the input/output model requires some criteria for

selecting the best model. Some statistical criteria are used for this selection.

Regarding the problem of the lack of measurements of important variables, estimation theory has introduced the Kalman filter, an optimal estimator capable of estimating all the states in the system. This optimal estimator, nevertheless, is based on models of the system. The use of Volterra series as a model to estimate states with the Kalman filter is then introduced. For this, a state space representation of the Volterra series is elaborated. The state space representation is the corner stone of modern control theory. Kalman filter uses the state space representation which has been classically represented by differential equations coming from fundamental physical laws [4], but here it is shown how an input/output based model can be represented in state space form and be coupled to Kalman recursions.

The applications of modeling based on Volterra series are made to cell recycle and batch fermentations for ethanol production using a recombinant *Escherichia coli*. Cell recycle is a way of getting high cell concentrations and operating at dilutions higher than the maximum specific growth rate. Here it is used as an experimental case study to test the performance of Volterra series modeling. Batch experiments are used to test the predicting capabilities of the Volterra series models. A model (Volterra series) obtained from one experiment is used to predict the same state variable in a similar experiment under the same operating conditions.

2 Input/Output Models and Estimation

A dynamical system can be described externally by specifying how the system responds to different stimuli without specifying the internal properties. Kalman called this the *external, input/output* or *empirical* definition of a dynamical system. One basic tool employed in the external description of a non-linear system is the so-called *Volterra series*, which has a history back to early this century.

For practical applications, the Volterra series in discrete form is as follows:

$$y(k) = \sum_{n=1}^{N} \sum_{i_1}^{\infty} \cdots \sum_{i_n}^{\infty} h(i_1, \ldots, i_n) u(k - i_1) \cdots u(k - i_n) \quad k = 0, 1, 2 \ldots . \tag{1}$$

The input signal $u(k)$ and output signal $y(k)$ are real sequences that are assumed to be zero for $k < 0$. The kernel $h(i_1, \ldots, i_n)$ is real and equal to zero if any argument is negative. The above system is stationary, causal and it degree-n homogeneous. This a polynomial of order N.

The determination of kernel values for an unknown system from a general input/output description is a linear problem. A simple Least Squares (LS) solution is sufficient to determine the best estimates of the kernels. One method that is not restricted to any kind of input, and that has shown good estimates is that of solving the LS problem with a QR factorization. Algorithms for solving this problem are found in several references [5, 2].

For a certain value of N and a specific memory M, there exists a Volterra series model that fits a certain set of data. The "best" model defined is that which represents reality as closely as possible in a simple manner and is reproducible in similar situations. Two concepts that will be helpful to choose the best model for a particular case are the *periodogram* and *Akaike's Information Criterion or AIC.*

The periodogram $I(\omega_j)$, a function of frequency ω_j, in an informal way, is defined for a set of n observations x_t in terms of the discrete Fourier transform by $I(\omega_j) = n^{-1} \left| \sum_{t=1}^{n} x_t e^{-itw_j} \right|^2$. The more non-zero spectral values in the range of frequency, the more likely the signal will be to excite the system. For some systems it will be possible to create a signal with proper

non-zero values in frequency, but in some cases such as batch experiments this may not be feasible as batch processes are basically "initial value" processes.

The problem of overfitting a model is addressed by the Akaike information criterion (AIC). This is a theoretical approach that can be applied to a variety of situations [1, 7, 2]. If ones plots AIC(q) *versus* q, the number of parameters to fit, the graph will, generally, show a minimum value, and the appropriate order of the model is determined by the value of q at which AIC(q) attains its minimum value. An autoregressive term represents the output at one lag in the past and used as an input, say, $y(k) = a_1 y(k-1)+$ Volterra series.

2.1 State Estimation

It is known that the problem of state estimation has been addressed typically by the use of the Kalman filter. Traditionally this filter is used along the state space representation based on a set of differential equations (i.e. a model based or internal approach) [4]. The Volterra series is a model representation obtained by input-output relationships, capable of representing complex systems, thus, it is suggested using this non-model based approach coupled to the filter. Before using the Kalman recursions, however, it is necessary to represent the Volterra series into state space form.

The state space form of any system is represented by:

$$\mathbf{X}_{t+1} = F_t \mathbf{X}_t + \mathbf{V}_t, \quad t = 1, 2, \ldots, \text{ and } \mathbf{Y}_t = G_t \mathbf{X}_t + \mathbf{W}_t, \quad t = 1, 2, \ldots. \tag{2}$$

The first equation is called the *state equation* and describes the evolution of the system at time t. The second equation is called the *observation equation*, in which a sequence of observations, $\mathbf{Y}_t$, which are obtained by applying a linear transformation to $\mathbf{X}_t$ and adding a random noise vector, $\mathbf{W}_t$, representing the errors involved in the measurement. More details on the assumptions of this representation can be obtained in other references [1, 10, 8, 7]. It is also clear from the definition that neither $\mathbf{X}_t$ nor $\mathbf{Y}_t$ is necessarily stationary.

For a second order Volterra series given by,

$$x(k) = h_0 + \sum_{i_1=1}^{M-1} h(i_1)u(k-i_1) + \sum_{i_1=1}^{M-1}\sum_{i_2=1}^{M-1} h(i_1, i_2)u(k-i_1)u(k-i_2). \tag{3}$$

Without loss of generality, the constant term h_0 can be included in $x(k)$. Now the state vector $\mathbf{X}_t$ of lower dimension (M x 1) is defined by $\mathbf{X}_t = \begin{bmatrix} x_{t-1} & u_{t-1} & u_{t-2} & \cdots & u_{t-(M-1)} \end{bmatrix}'$. In reality, the input process U_t is non-stationarity and because of this, certain manipulations can be done to improve the representation according to Brockwell and Davis [1].

Once the state space has been defined, the aim will be that of finding the best estimates of the state vector $\mathbf{X}_t$ in terms of the observations $\mathbf{Y}_1, \mathbf{Y}_2, \ldots$. More specifically, the problem will be solved when the estimates of $\mathbf{X}_t$ in terms of $\mathbf{Y}_0, \ldots, \mathbf{Y}_t$ are found, i.e., solving the *filtering problem.*

The best estimator $\mathbf{X}_{t|k}$ of $\mathbf{X}_t$ is the random vector whose ith component, $i = 1, \cdots, v$, is the best linear estimator of X_{ti} in terms of all the components of $Y_0, Y_1, \cdots, Y_k$ (specifically, for the filtering problem $k = t$). In particular, $\hat{\mathbf{X}} = \mathbf{X}_{t|t-1}$. The solution algorithm can be found in several references.

One of the assumptions in the Kalman recursions is that the random vector $\mathbf{U}_t$ must be uncorrelated, i.e., white noise. If this vector is not white noise, a proper transformation to the state equation can be made, in particular to the dynamic matrix F_t. The idea is to model the

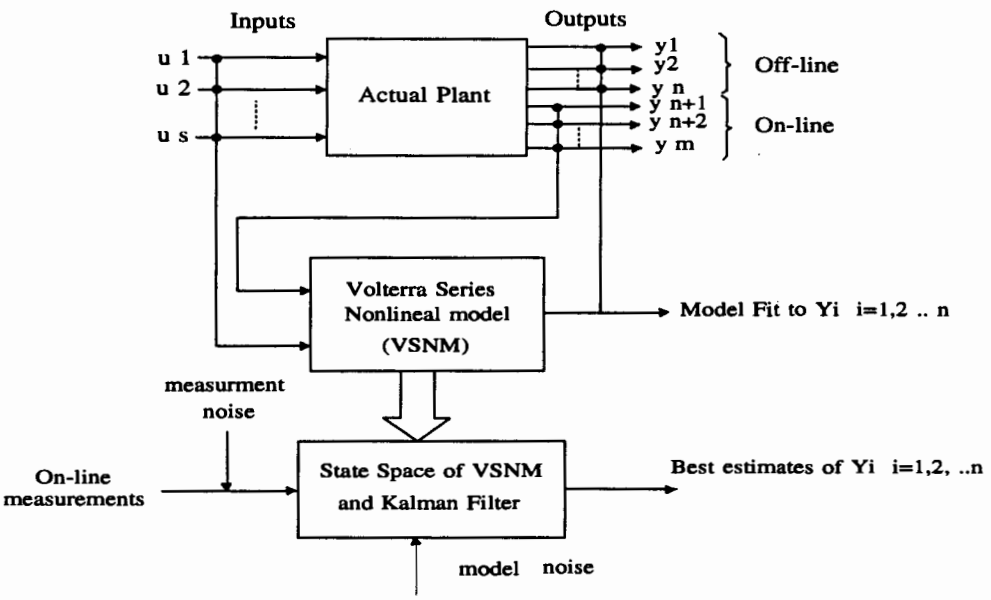


Figure 1: Block diagram of overall objective

error vector $\mathbf{V}_t$ and include this new model as an additional row in the dynamic matrix. This will imply that the dimension of the state vector is incremented by one.

For the sake of simplicity, an Autoregressive (AR(p)) model is used for modeling the error. In practice, this model is easy to estimate and usually a small number of p is sufficient for representing real cases. Supposing that an AR(1) model is the best model, the new dynamic matrix F_t of order $((M+2)\text{ x }(M+2))$ has the following form,

$$F_t = \begin{bmatrix} 0 & 0 & \frac{\partial x_t}{\partial u_{t-1}} & \cdots & \frac{\partial x}{\partial u_{t-(M-2)}} & \phi^* \\ 0 & \phi_1 & 0 & \cdots & & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ & & 1 & 0 & & \vdots \\ & & & \ddots & & 0 \\ 0 & \cdots & & & 0 & \phi^* \end{bmatrix}, \tag{4}$$

where the last row represents the parameters of the AR(1) model.

The modified state vector $\mathbf{X}_t$ is given by $\mathbf{X}_t = \begin{bmatrix} x_{t-1} & X_t & u_{t-1} & \cdots & u_{t-(M-2)} & X_t^* \end{bmatrix}^t$ where X_t is the stationary random variable needed to represent the non-stationary input variable U_t, and X_t^* is the stationary random variable for describing the colored nature of the colored model error. If an autoregressive term is considered, the constant a_1 is used in the position (1,1) of the matrix F_t.

At real time, only the on-line measurements and the value(s) of the manipulated variable(s) (if any) are available and it is desired to develop a scheme that allows one to know in real time the value of the rest (and more important) state variables, since the latter is the most important indicator of the process and the ones used for optimization and control techniques. Therefore, the first step will be to fit a nonlinear model using the Volterra series. To do this, some previous experiments must be run and analyzed thoroughly in order to provide as much information as possible (on-line and off-line measurements). This will help construct a more reliable and

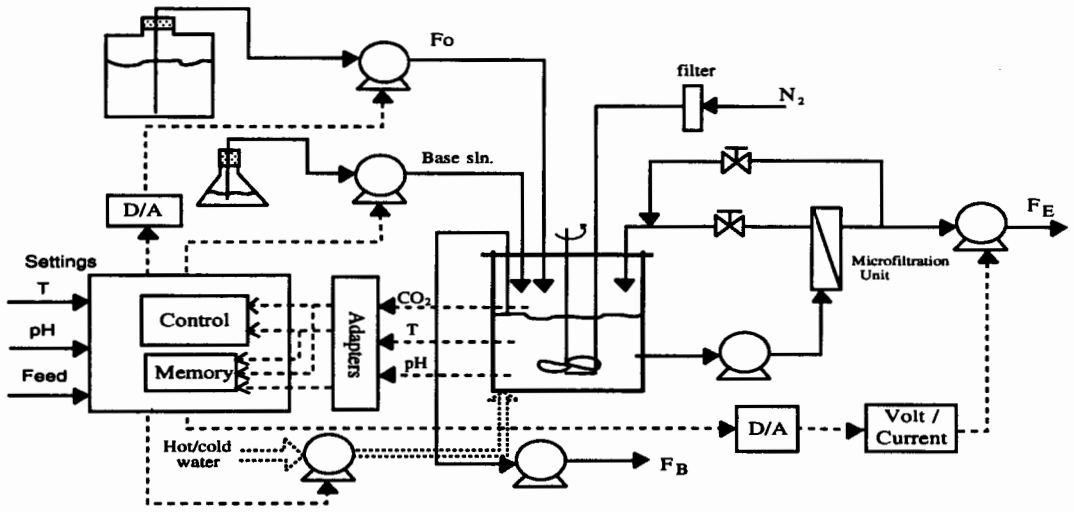


Figure 2: Flow diagram of experimental cell recycle system.

accurate model.

However, the model obtained thus far will not be able to provide the best estimates on-line for the important states, especially if in reality there are more facts to consider. Among them are the measurement and model noises. The Kalman filter is used to minimize the effects of these. The overall process can be simplified schematically as in Figure 1.

3 Experimental Systems

The Volterra series was applied to a fermentation process, a system commonly known to show modeling difficulties because of its nonlinear and time varying nature. One goal in every fermentation is to increase the total productivity in one product or the biomass. In continuous systems, one desirable feature is obtaining as high amount of cells as possible. Cell recycle is one way to accomplish this. Cell recycle is a process where microbial cells are recovered from the effluent of the fermentor and recycled back into the fermentor. Figure 2 shows the experimental system built. The recirculation R is defined as $R = \frac{F_E}{F_E+F_B}$. If $F_B = 0$, i.e. $R = 1$, then no steady state is possible because these condition represent a fed-batch operation. For this type of chemostat $\mu = (1 - R)D \leq D$.

4 Materials and Methods

One cell recycle fermentation was tested. The strain *Escherichia coli* ATCC 11303 containing plasmid PLOI297 in anaerobic condition to convert xylose to ethanol was used. The bacteria utilized was grown in a Luria broth [6]. The initial concentration of xylose was 40 g/l. The operating volume was 2 l. A batch fermentation was the first phase until a value of 20-30 g/l of xylose was reached; then the cell recycle took place. The feed concentration used was 60 g/l of xylose and the proper concentrations of additional nutrients were according to the

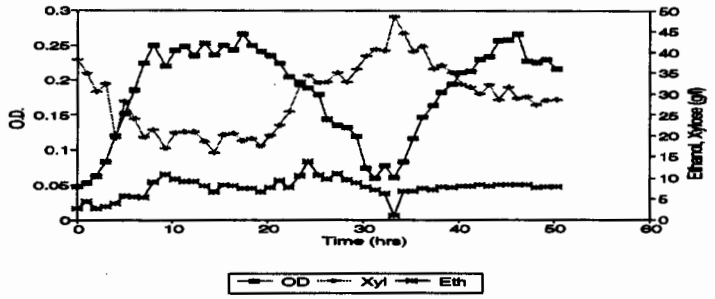


Figure 3: Off-line measurements in cell recycle.

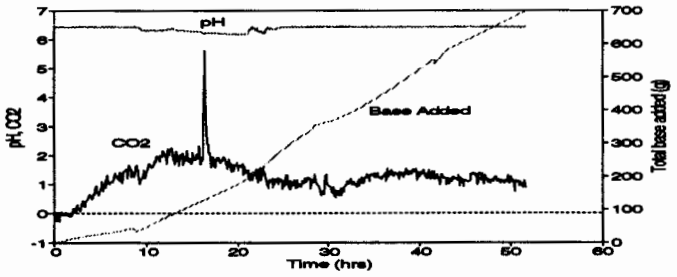


Figure 4: Some on-line measurements in cell recycle fermentation.

Luria broth. The operating conditions are 34° C, pH=6.5 and 150 rpm. The dilution rate was initially set to 0.2 hr^{-1} and a recirculation rate of 0.5.

The system was instrumented based on a PC microcomputer equipped with two D/A cards to activate the feed and effluent pumps. Temperature, pH and CO_2 evolution on-line were stored on-line. The fermentor used was a Bench-Top New Brunswick fermentor with a total capacity of 7 l and a MF-114 module.

The separation of the biomass from the broth was accomplished by a microfiltration unit with 0.2 micron as pore size. The CO_2 gas composition in the air outlet of the fermentor was measured with a gas analyzer. The pH in the fermentor was kept constant adding an alkali 4N solution of NaOH. The amount of base added is recorded for the solution over a balance that sends digital signals to the card. Nitrogen must be continuously sparged into the fermentor making the system anaerobic.

To obtain the most out of the dynamic behavior of the system, various conditions of dilution rate were used and a fixed recirculations. The literature shows common recirculations of 0.5 is common in bacterial fermentations [3]. A small amount is injected in the HPLC from which the concentrations of xylose, ethanol and various acids are measured.

Figures 3 and 4 show some off-line and on-line measurements obtained from a cell recycle fermentation. In the on-line graph, more than 600 points are stored. Meanwhile, the off-line analysis shows only 50 samples, each taken every hour. This indicates the necessity of estimators: reducing the amount of labor and measurement errors, and capability of offering the state values as often as the on-line measurements.

Actual recirculation ratios are estimated to be in the range of 0.2 to 0.35. N represents the number of points used to fit the model; M represents the memory and the AIC is the goodness of the model. The results shown here are focused on modeling mainly the OD (or cell mass). The measurement of cell mass or OD as an output variable could be a reliable control variable

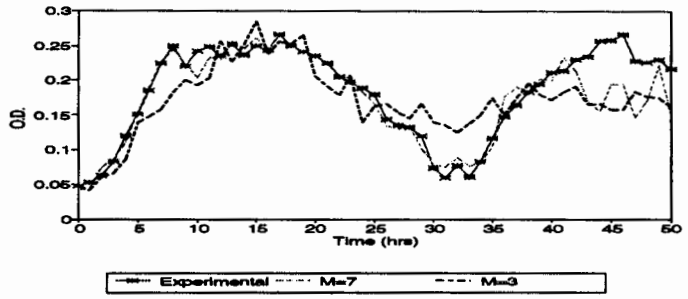


Figure 5: Second order Volterra series model of OD in cell recycle using CO_2 as input with N=40.

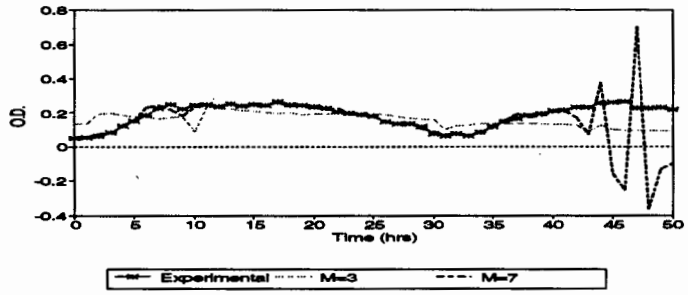


Figure 6: Modeling of OD in cell recycle using the total base-added as input.

since tetracycline is added to avoid plasmid segregation instability. The important conclusions drawn from this example of modeling are valid for any desired output.

Figure 5 shows the best one input-one output modeling when the input is the CO_2 evolution and the output is the OD. The minimum AIC (-240.18) for N=40, corresponds to the model with M=3.

The Volterra series does a good job when using it with the same number of points as used for the fitting. It is not a good tool for evaluating (estimating) the values at points in the prediction range ($t > 40$). It was found that the higher the memory, the better the model fit inside the range and the prediction is worse. Then two objectives be balanced. This can be done with the aid of the AIC. This parameter shows that the minimum AIC indicates the best model and may not necessarily indicate the smallest overall sum of errors. The AIC gives a model which behaves better in the prediction area. Use of high memories usually results in erroneous predictions. The latter is due to the overfitting of parameters in the model.

Base-added is a less appropriate signal as an input since the "estimates" are even worse for points outside the range (see Figure 6). A convincing way of showing why CO_2 is more appropriate than base-added for building models is by analyzing the power spectrum of the signals (looking at the *periodogram*). Figure 7 shows the periodogram of CO_2. The corresponding plot for base added shows practically no weight at any frequency. In other words, this signal does not have energy enough to excite the system. Whereas, the CO_2 periodogram shows bigger components in more varied frequencies (enough to excite the system in some form).

The CO_2 series is converted to a stationary process with a differencing operation and modeled as an autoregressive model, as $X_t = \phi_1 X_{t-1} + Z'_t$, $Z'_t \sim WN(0, \sigma^2)$, with $\phi_1 = -0.514435$ and $\sigma^2 = 0.2608$. X_t is a stationary variable which is the result of applying differencing once to the input CO_2, u_t. A similar procedure is done for modeling the error between the model

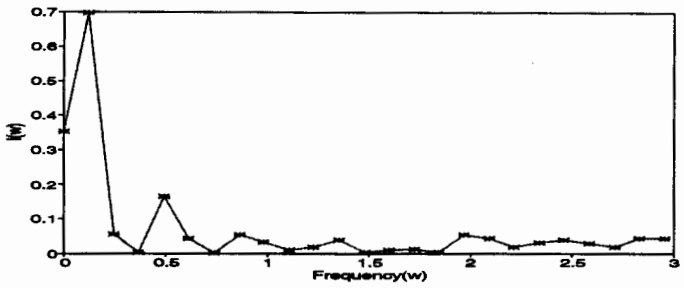


Figure 7: Periodogram of CO_2 in cell recycle experiment.

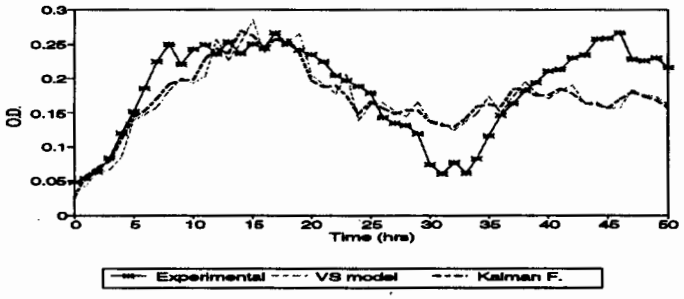


Figure 8: Estimation of OD using Kalman filter based on Volterra series model.

and the experimental results. From the autocovariance of the measurement error, it has been shown that the random part of the measurement is practically white noise [9].

All parameters, needed for the Kalman filter, including initial values, are obtained statistically from measurements and model errors. No value is tuned by trial and error. Figure 8 shows the modeling and estimation of OD using only Volterra series, while Figure 9 shows the improvement in prediction when the autoregressive term is considered.

5 Batch Fermentations

The purpose of this section is to show to the applicability of one Volterra series model from one experiment to estimate the state(s) of another similar experiment. One idea is that of defining how similar two experiments can be, i.e., the model from one experiment can help

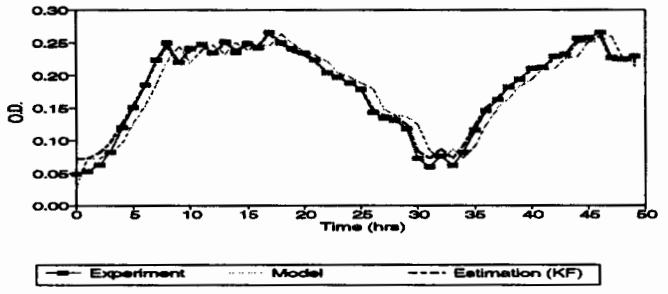


Figure 9: Estimation of OD using Kalman filter based on Volterra series model plus an autoregressive term.

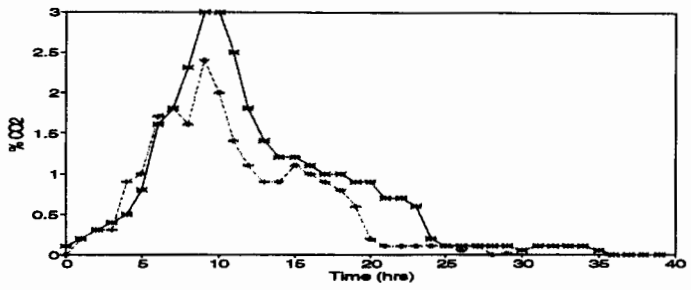


Figure 10: CO_2 concentrations for two batch fermentations.

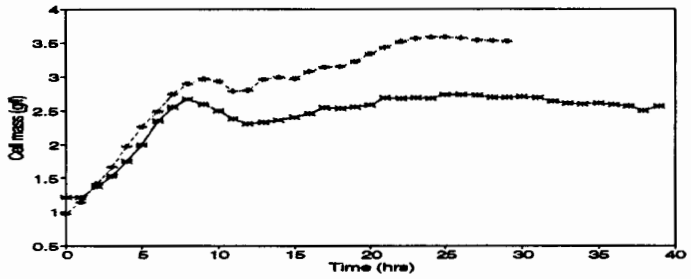


Figure 11: Cell mass concentrations for two batch fermentations.

decide whether another experiment is reasonably similar to a characteristic one whose model has been obtained. If one experiment does deviate from the model, it can be said that the experiment is not characteristic and something is erroneous.

To address this issue two batch experiments were studied. Both experiments were performed using the same strain, media and conditions in the same fermentor as those used in the cell recycle operation.

Results of two batch fermentations are presented. Figures 10 and 11 show the the on-line CO_2 measurements that will be used as inputs and the cell mass concentrations to be used as outputs, respectively.

The experiment that was run for 40 hours is selected for modeling because of the larger number of points. This is referred to as experiment **A**. Its results are easily distinguished from the other experiments due to this fact. All models were carried out considering all available points, i.e., no prediction is performed.

Considering the autoregressive model, according to the AIC, the model that minimizes this criteria is when M=2. Figure 12 shows how experiment **B** is estimated from the model of experiment **A**. Studying Figure 11, it may be concluded that the experiments are not similar (one is not representative), but mathematically, the CO_2 to cell mass relationship is similar and thus, these experiments are alike. Some statistical criteria are needed to define the boundaries of confidence for determining if an experiment is acceptable.

6 Conclusions

It is shown here that the Volterra series model can represent a highly nonlinear fermentation process. The proposed method in predicting states of cell recycle fermentations (continuous)

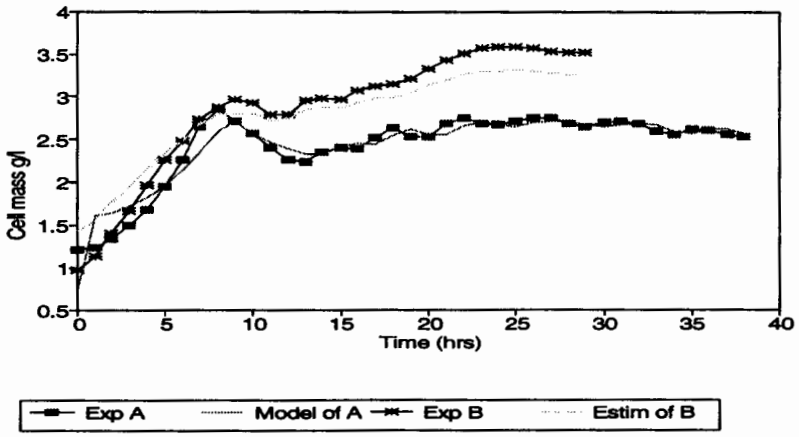


Figure 12: Estimation of cell mass concentration in batch experiment **B** using the best model from **A** coupled to Kalman filter.

is attractive for optimization and control techniques; whereas in batch fermentations, can be considered as an additional tool in performance monitoring.

References

[1] Brockwell, P. and Davis, R.A. *Time Series : Theory and Methods.* Springer-Verlag, second edition, 1991.

[2] Chen, S., Billings, S.A. and Luo, W. Orthogonal least squares methods and their application to nonlinear system identification. *Int. J. Control*, 50(5):1873–1896, 1989.

[3] De Boks, P.A. and Van Eyberggen, G.C. Continuous ethanol production using cell recycle with a settler. *Biotech. Lett.*, 3:577, 1981.

[4] Friedland, B. *Control Systems Design.* McGraw-Hill, 1987.

[5] Golub, G.H. and Van Loan, Ch. *Matrix Computations.* Johns Hopkins University Press., second edition, 1989.

[6] Hilaly, A.K. *Structured mathematical modeling and optimization of xylose fermentation by a recombinant Escherichia coli.* PhD thesis, Colorado State University, 1992.

[7] Priestley, M.B. *Spectral Analysis and Time Series.* Academic Press, 1981.

[8] Subba Rao. *An Introduction to Bispectral Analysis and Bilinear Time Series Models.* Springer Lecture Notes, 24, second edition, 1984.

[9] Saucedo, V.M. Modeling of recombinant fermentations using volterra series. Master's thesis, Colorado State University, March 1993.

[10] Tong, H. *Non-linear Time Series: A Dynamical Systems Approach.* Oxford University Press., 1990.

DEADEND MICROFILTRATION OF BOVINE SERUM ALBUMIN SUSPENSION THROUGH YEAST CAKE LAYERS AND ASYMMETRIC POLYMERIC MEMBRANES

NAVEEN ARORA and **ROBERT H. DAVIS**
Department of Chemical Engineering
University of Colorado, Boulder, CO 80309-0424.

Abstract

The present work involves deadend filtration of bovine serum albumin in a phosphate buffer through a cake layer deposited on an asymmetric, polysulfone microfiltration membrane. Yeast cake layers of varying thicknesses are preformed in two ways -- by gravity settling prior to the onset of filtration and by vacuum filtration. The focus is on studying the role of the cake layer of rejected cells as a secondary membrane for reducing protein transmission and affecting permeate flux. Due to the micro-channeling effects, higher fluxes are observed for filtration through a cake layer preformed by vacuum filtration, as compared to the corresponding values for a cake layer of equal dry mass preformed by gravity settling. The steady state flux is higher in the presence of a thin cake layer than in the absence of a cake layer. This is thought to result from the yeast cake serving as a filter aid to remove aggregated proteins which would otherwise foul the membrane. The protein transmission studies reveal three types of fouling -- membrane fouling due to the internal pore deposition of protein, protein adsorption through the entire cake depth, and surface filtration leading to protein rejection at the cake surface.

INTRODUCTION

An ideal membrane process for recovering proteins from a suspension containing cells or cell debris is characterized by a high steady flux and 100% protein transmission. Flux is defined as the volume of permeate collected per unit time per unit membrane area, and protein transmission is defined as the concentration of protein in the permeate divided by that in the feed. Both flux and protein transmission decline over time as proteins adsorb onto the membrane or as the surface cake deposit of rejected biomass is rendered gradually less permeable, for example by excretion or retention of the macromolecules in the feed stream. Studies of protein-membrane interactions and protein adsorption on membranes have been carried out with both polymeric and inorganic microfiltration membranes (Tracey and Davis (1993), Fane *et.al.* (1992), Bowen and Hughes (1990), Le and Gollan (1989), Truskey *et.al.* (1987)). However, quantitative information is lacking on the effects of biomass/cake formation on the filtrate flux and the protein transmission during membrane filtration.

The purpose of the present paper is to report the details of a study of the microfiltration of bovine serum albumin (BSA) suspension in a phosphate buffer through an asymmetric polysulfone membrane with 0.2 μm nominal pore size in the presence of a yeast cake layer deposited on the membrane prior to the onset of filtration. Such an arrangement provides a situation analogous to that normally occurring when the microfiltration membranes are used to separate proteins from cells or cell-debris. The focus is on investigating the role of the cake layer of rejected cells as a secondary membrane for reducing protein transmission and affecting the filtrate flux profiles.

MATERIALS AND METHODS

The equipment (manufactured by Millipore) used for BSA microfiltration experiments consists of a borosilicate glass funnel and a ground glass base with coarse-frit glass support for the membrane. The ground glass base is equipped with a neoprene stopper which fits tightly into a 1000 ml collection flask. A rubber gasket is used to provide an efficient contact between the funnel and the filter base surfaces. The microfilters used in these filtrations are modified hydrophilic, low protein binding, asymmetric, polysulfone membranes (manufactured by Gelman Sciences) in the form of circular discs of diameter 47 mm, nominal pore size 0.2 μm, and thickness 150 μm. These modified polysulfone membranes, coated with a layer of hydroxy ethyl methacrylate, offer higher flow rates, lower extractables and greater strength than competitive cellulose acetate or cellulose nitrate membranes.

The collection flask is subjected to a vacuum of 18 in Hg (8.8 psi) in order to effect the permeate flow through the microfiltration membrane. An on-line vacuum gauge (Fisher Scientific), having a range of 0 to 30 in Hg, is used to measure pressure in these experiments. The suspension used for yeast cake formation is *Saccharomyces cerevisiae* (Fleischmann's active dry yeast, reconstituted) suspended in a deionized water having a resistivity of 18.2 megaohm-cm. Prior to being used for cake formation, the suspension containing a known dry mass of yeast cells is subjected to three runs of centrifugation in a Beckman GPR centrifuge for washing. Each centrifugation run is carried out at 2000 rpm, for a period of 5 minutes. The supernatant containing the yeast cell fragments, and the surface active proteins separated from the yeast cells, is discarded at the end of each centrifugation run.

Phosphate buffer solution (PBS) serves as the suspending medium for BSA and is also used to determine the membrane resistance in these experiments. PBS is prepared by dissolving 8.0 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, and 0.24 g KH_2PO_4 in a final volume of 1000 ml deionized water. Bovine serum albumin solution is prepared by suspending BSA crystals (Sigma Chemical Co., grade A6793, fractionated by heat shock, 98-99% albumin) in the buffer solution having a pH of 7.4. BSA is a well characterized protein of molecular weight 67,000, size 11.6 x 2.7 x 2.7 nm, and with the isoelectric point at pH 4.7-4.9 (Bowen and Gan, 1991). The concentration of BSA solution used in all filtrations is 0.1 mg/ml.

Prior to the yeast cake layer experiments, the membrane resistance, R_m, is determined by measuring the flow of the buffer solution through the membrane, as a function of time. The flow rate is measured by recording the time required for a known volume of feed solution to pass through the microporous membrane. Similar flux measurements are also conducted for BSA microfiltration through the membrane in the absence of yeast cake layer. Each yeast cake layer experiment consists of filtering BSA in the phosphate buffer through the yeast cake layer and the microporous membrane for a certain period of time (typically 5 hours) until the flux becomes nearly steady. Yeast cake layers of varying thicknesses are preformed in two ways -- by gravity settling of the yeast cell suspension prior to the onset of filtration, and by vacuum filtration prior to the start of BSA filtration. The flux measurements are followed by protein assays, performed to determine the ratio of absorbance of permeate samples to that of feed samples, at different times during the experiment, thus obtaining the protein transmission variation with time or with the permeate volume collected. The absorbance measurements, for calculating BSA concentration

in the feed and the filtrate, are conducted in a Beckman DU-50 spectrophotometer at a wavelength of 280 nm.

RESULTS AND DISCUSSION

Figure 1 shows the results of the experiment in which phosphate buffer solution (PBS) is filtered through the microporous membrane. Each data point represents the arithmetic mean of the corresponding values obtained from the three runs of the experiment, and the error bars represent 90% confidence intervals. Figure 2 highlights the time-dependent membrane resistance calculated, based on the permeate flux data from Fig. 1, using Darcy's law for membrane filtration in the absence of a cake layer. Since, the membrane resistance is an intrinsic property of the membrane, and PBS causes no fouling, the value of R_m for PBS filtration is expected to remain constant with time. However, owing to the membrane compaction upon the imposition of the pressure drop, a rapid but modest increase in R_m is observed in the initial few minutes of PBS filtration (Fig. 2). For the employed polysulfone membranes, the steady state value of R_m (with 90% confidence) is found to be $(4.2 \pm 0.1) \times 10^8$ cm^{-1}.

The permeation rate for BSA filtration through the membrane in the absence of a cake layer is found to decrease continuously with time, over a period of about 3 hours (Fig. 3). However, the protein transmission levels off to unity in approximately 6 min after the onset of filtration (Fig. 4). Such a continual decrease in filtrate flux with time, under constant applied pressure, is much greater than can be explained by protein adsorption, which is generally limited to monolayer coverage. Such decreases have been interpreted in terms of protein deposition resulting from shear-induced distortion of the molecules on passage through the micropores. Circular Dichroism (CD) and Fourier Transformed Infrared Spectroscopy (FTIS) have shown that changes in secondary structure of protein can occur on passage through microfiltration membranes (Truskey *et.al.*, 1987). The decrease in flux with time can be quantitatively described as a decrease in pore volume or number due to internal fouling of the membrane caused by protein deposition at the walls or mouths of the pores. Tracey and Davis (1993) have found that during internal fouling, rather than accumulating uniformly along the pore walls, proteins accumulate near the entrance of the pores and on the membrane surface. Kelly *et.al.* (1993) have explained the flux decline in terms of denatured protein aggregates fouling the membrane and serving as initiation sites for the continued deposition of bulk (native) protein. The minimum level of protein transmission for BSA filtration through the membrane in the absence of a cake layer is observed approximately 2 min after the onset of the experiment (Fig. 4). This time period possibly accounts for the growth of denatured protein agglomerates which promote further protein adsorption of bulk (native) protein on the pores of the membrane. The subsequent increase in protein transmission can be attributed to near saturation coverage (which corresponds to deposition of protein monolayer) of the internal structure of the membrane.

Figures 5 and 6 depict the time-dependent filtrate flux profiles obtained for BSA microfiltration through the polymeric membranes in the presence of yeast cake layers of varying thicknesses preformed by vacuum filtration and by gravity settling of the yeast cell suspension, prior to the onset of filtration. Cake layers of 0.2 g, 0.4 g, and 0.8 g (dry mass) yeast are formed in these experiments. These yeast dry masses lead to the formation of 0.014 g/cm^2, 0.028 g/cm^2, and 0.056 g/cm^2 yeast cake layers, respectively, based on the external surface area of the

deposited cake layer. The corresponding cake thicknesses, calculated using the appropriate conversion factor correlating the yeast cake dry mass with the cake volume (Ju and Ho, 1987), are found to be 0.051 cm, 0.102 cm and 0.204 cm, respectively.

As shown in figures 5 and 6, lower long-time fluxes are observed for filtration through a cake layer preformed by gravity settling, as compared to the corresponding values for a cake layer of equal dry mass preformed by vacuum filtration, prior to the start of BSA microfiltration. This can be attributed to the fact that the vacuum filtration cake layer, which is formed under suction, has micropores incorporated in its structure which provide low resistance flow channels, thereby leading to higher permeation rates than those for the cake layer preformed by overnight gravity settling of the yeast cell suspension. Also, the fact that the initial flux values associated with a gravity cake are found to be approximately the same or slightly higher than those for a vacuum cake of the same mass leads to the conclusion that the former is less compact to begin with. Thus, initially, the less compact cake structure of the gravity settled cake layer counterbalances the micro-channeling associated with the vacuum filtration cake layer, in terms of affecting the filtrate flux. With time, however, a pressure drop imposed across the gravity cake causes it to compact, so that the flux declines to a lower value than that of the precompacted vacuum cake. Comparison of figures 5 and 6 to figure 3 leads to the conclusion that the long-time flux for BSA filtration in the presence of 0.014 g/cm^2 and 0.028 g/cm^2 (0.2 g and 0.4 g dry mass, respectively) yeast cake layers is greater than the long-time flux in the absence of one, thus illustrating the role of the cake layer as a pre-filter for the bovine serum albumin solution, besides providing the cake resistance for BSA filtration. It is hypothesized that the yeast cake removes aggregated or denatured proteins which would otherwise foul the membrane. The cake resistance offered by the 0.056 g/cm^2 (or 0.8 g dry mass) yeast cake layer, however, predominates its role as a pre-filter for BSA solution, thereby leading to no improvement in the long-time flux for BSA microfiltration.

Figures 7 and 8 illustrate the protein transmission profiles obtained for BSA microfiltration in the presence of the vacuum filtered and the gravity preformed yeast cake layers of varying thicknesses. Since, the microfiltration membranes used in these filtrations exhibit 100% protein transmission in the absence of a cake layer, except during the first few minutes of filtration, the low levels of protein transmission observed initially in the preformed cake layer experiments can be attributed to protein adsorption through the entire cake depth. Gradually, the protein transmission levels off to a constant value of approximately 95%, independent of the cake thickness (or dry mass) and the cake formation pattern. As anticipated, the rise in protein transmission is found to be more gradual for a thicker yeast cake as compared to that observed for a thinner cake layer. The small offset from 100% transmission, observed even at large times, is possibly due to protein rejection at the cake surface. This phenomenon of depth filtration followed by surface filtration is observed for the three thicknesses of both the vacuum filtered and the gravity settled yeast cakes.

Figures 9 and 10 give the plots of the mass of BSA rejected (excluding protein rejection due to surface filtration) per unit external surface area of the cake versus the cake dry mass per unit external surface area. The mass of BSA rejected is calculated based on the data obtained from the protein transmission measurements conducted for different thicknesses of yeast cake layers. The mass of BSA rejected (due to protein adsorption through the entire cake depth) appears to

have varied linearly with the cake dry mass, for both the gravity and the vacuum preformed cakes, within the experimental error limits.

CONCLUSIONS

Experiments performed to study the deadend microfiltration of BSA through asymmetric polysulfone membranes in the presence of the yeast cake layers confirm that the permeate flux profiles are significantly influenced by the cake formation pattern. Owing to the micro-channeling effects associated with the vacuum filtration cake layer, and also to the compressibility of the gravity settled cake layer, lower long-time fluxes are observed for filtration through a cake layer preformed by gravity settling as compared to the corresponding values for a cake layer of equal dry mass preformed by vacuum filtration. The fact that the long-time flux in the presence of a thin cake layer is found to be greater than that in the absence of a cake layer leads to the inference that the cake layer acts as a pre-filter for the protein solution. The protein transmission studies reveal three types of fouling -- membrane fouling due to the internal pore deposition of protein, protein adsorption through the entire cake depth, and protein rejection at the cake surface. These protein losses may have an important influence on the overall yield of a small volume, high value, protein product of interest like BSA.

Acknowledgements

This research has been supported by Grant CTS 9107703 from the National Science Foundation and the University of Colorado's Summer Minority Access to Research Training (SMART) Program. The authors would like to acknowledge Paul Suding and Larry Gallegos for help in the experiments.

References

* Bowen, W.R. and Hughes, D.T. (1990). Properties of microfiltration membranes. Part 2. Adsorption of bovine serum albumin at aluminum oxide membranes. *J. Membrane Sci.*, **51** : 189-200.
* Fane, A.G., Kim, K.J., Hodgson, P.H., Leslie, G., Fell, C.J.D., Franken, A.C.M., Chen, V., Liew, K.H. (1992). Strategies to minimise fouling in the membrane processing of biofluids. In P.W. Todd, S.K. Sikdar, M. Bier (eds.), ACS Conference Proceedings Series, *Proceedings of Frontiers in Bioprocessing II*, pp 304-320.
* Ju, L.K. and Ho, C.S. (1987). Correlation of cell volume fractions with cell concentrations in fermentation media. *Biotech. Bioeng.*, **32** : 95-99.
* Kelly, S.T., Opong, W.S. and Zydney, A.L. (1993). The influence of protein aggregates on the fouling of microfiltration membranes during stirred cell filtration. *J. Memb. Sci.* (in press).
* Le, M.S. and Gollan, K.L. (1989). Fouling of microporous membranes in biological applications. *J. Membrane Sci.*, **40** : 231-242.
* Tracey, E.M. and Davis, R.H. (1993). Protein fouling of polycarbonate track-etched microfiltration membranes. In W.W. Leung (ed.), Proceedings of the 6th AFS Annual Meeting and Exposition, *Advances in Filtration and Separation Technology*, **7** : 427-430.
* Truskey, G.A., Gabler, R., Dileo, A., Manter, T. (1987). The effect of membrane filtration upon protein conformation. *J. Parenter. Sci. Technol.*, **41** : 180-193.

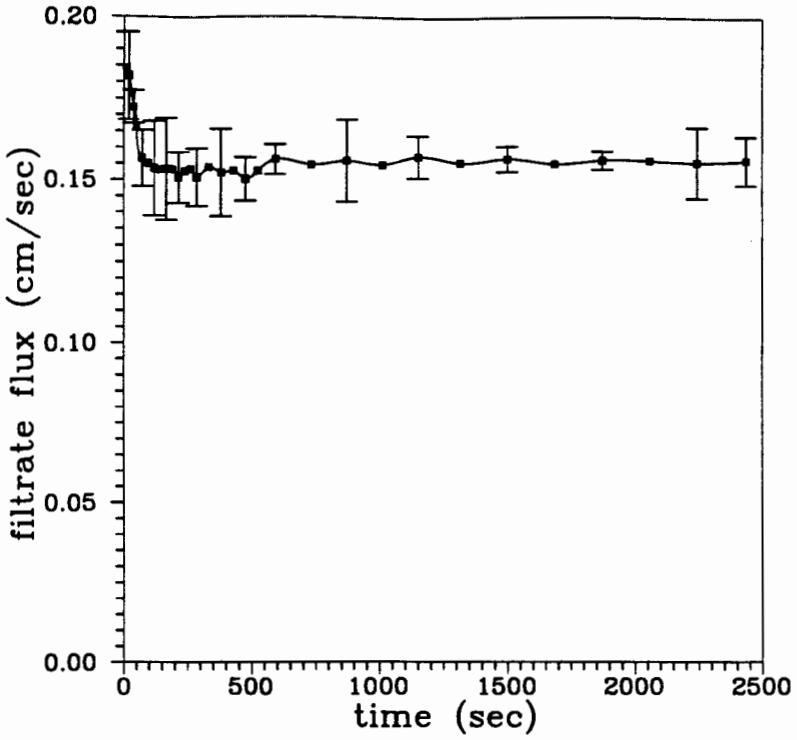


Figure 1: Variation with time of the permeation rate of phosphate buffer solution (PBS) through 0.2 μm, modified hydrophilic, asymmetric, polysulfone membrane. The error bars represent 90% confidence intervals for three experiments.

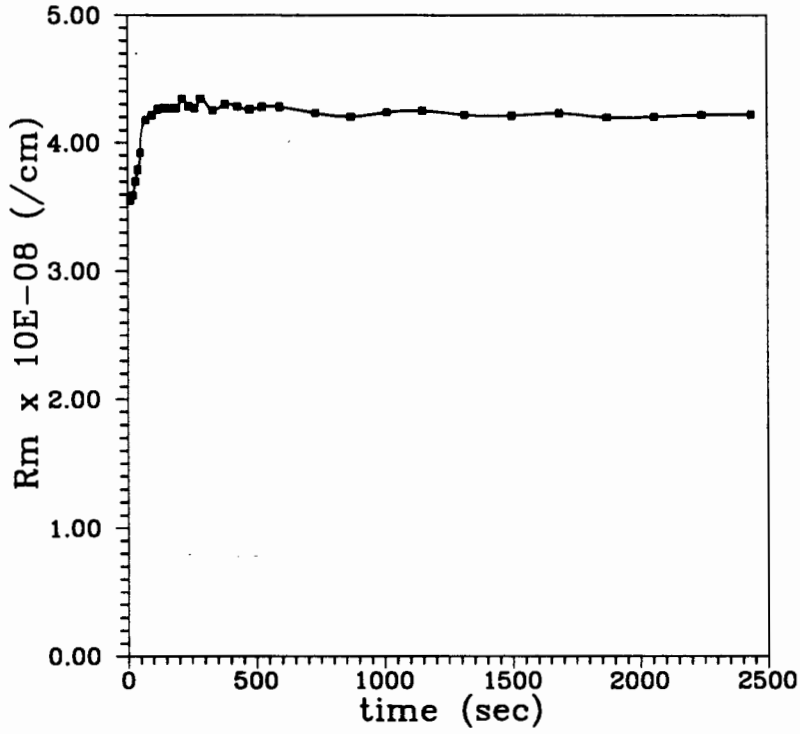


Figure 2: Time profile of the membrane resistance for PBS filtration, calculated based on the flux variation depicted in figure 1.

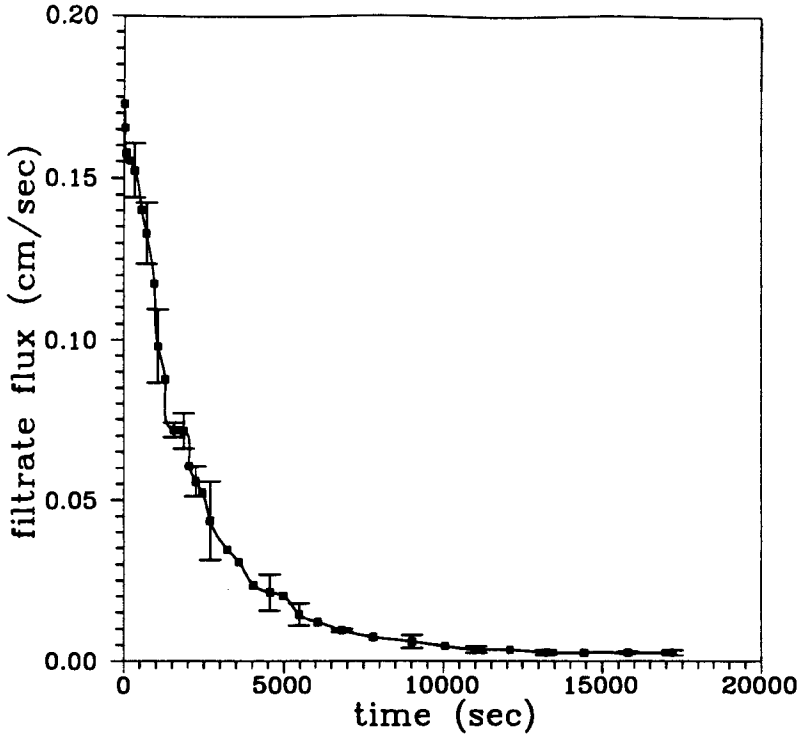


Figure 3: Variation with time of the permeation rate of 0.1 mg/ml BSA solution through 0.2 μm polysulfone membrane (in the absence of a cake layer). The error bars represent 90% confidence intervals for three experiments.

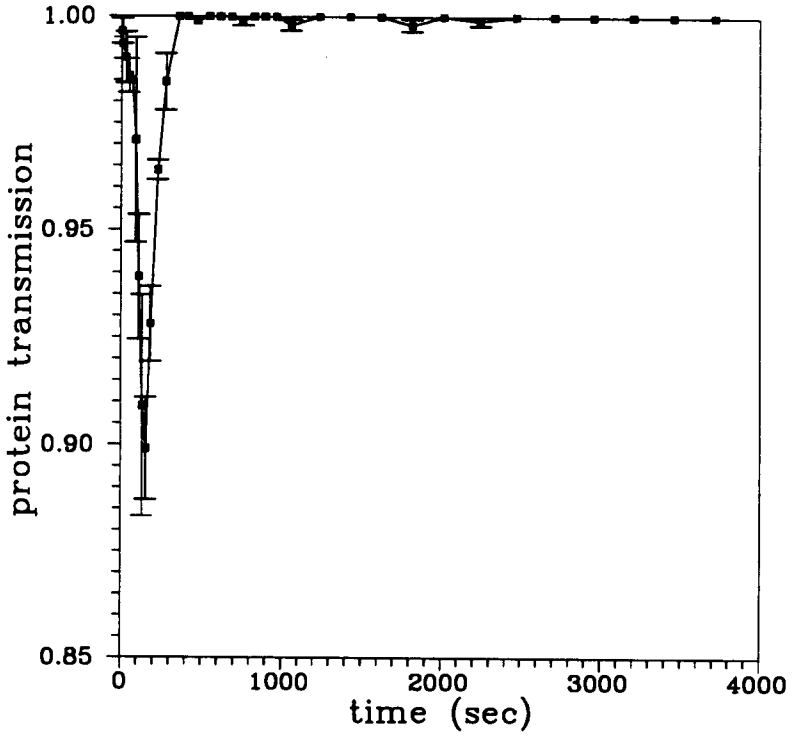


Figure 4: Protein transmission as a function of time, for 0.1 mg/ml BSA solution microfiltration through the polymeric membrane (in the absence of a cake layer). The error bars represent the 90% confidence intervals for three experiments.

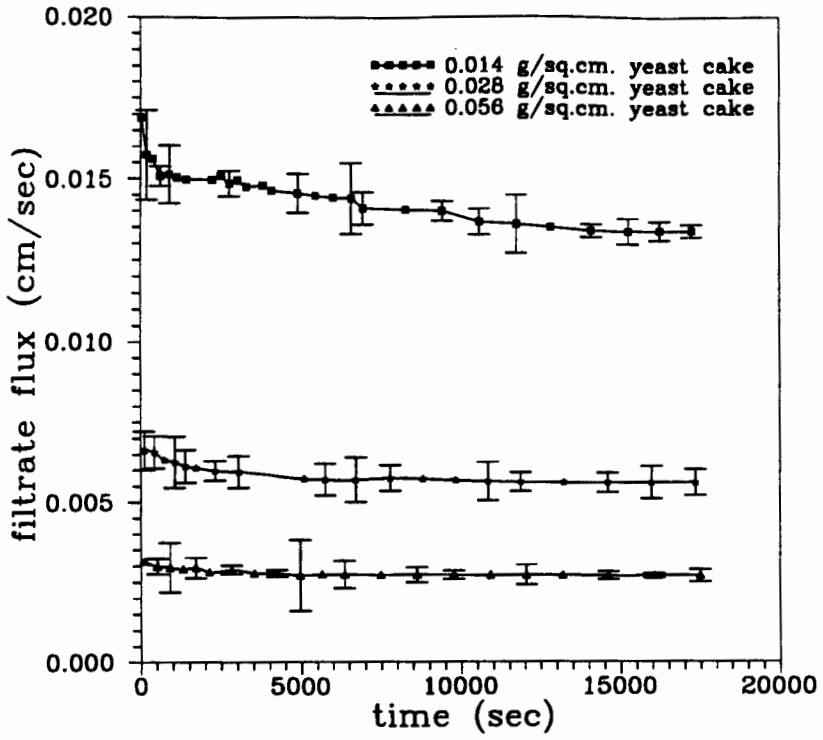


Figure 5: Filtrate flux variation with time, for 0.1 mg/ml BSA solution filtration through 0.2 μm polysulfone membrane in the presence of different thicknesses of vacuum filtered yeast cake layers. The error bars represent 90% confidence intervals for three experiments.

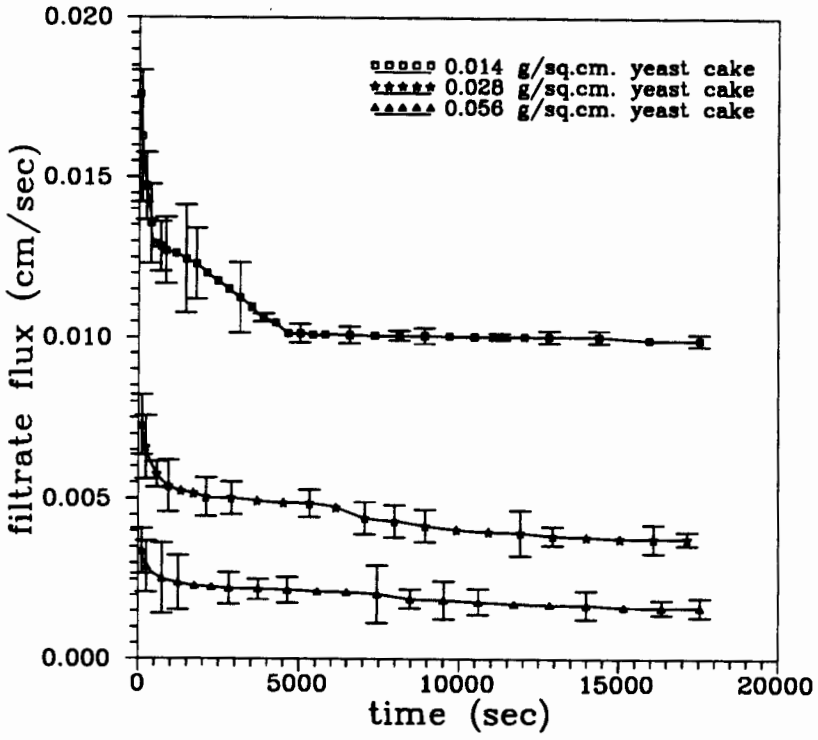


Figure 6: Permeate flux variation with time, for 0.1 mg/ml BSA solution filtration through the polymeric membrane in the presence of different thicknesses of gravity preformed yeast cake layers. The error bars represent 90% confidence intervals for three experiments.

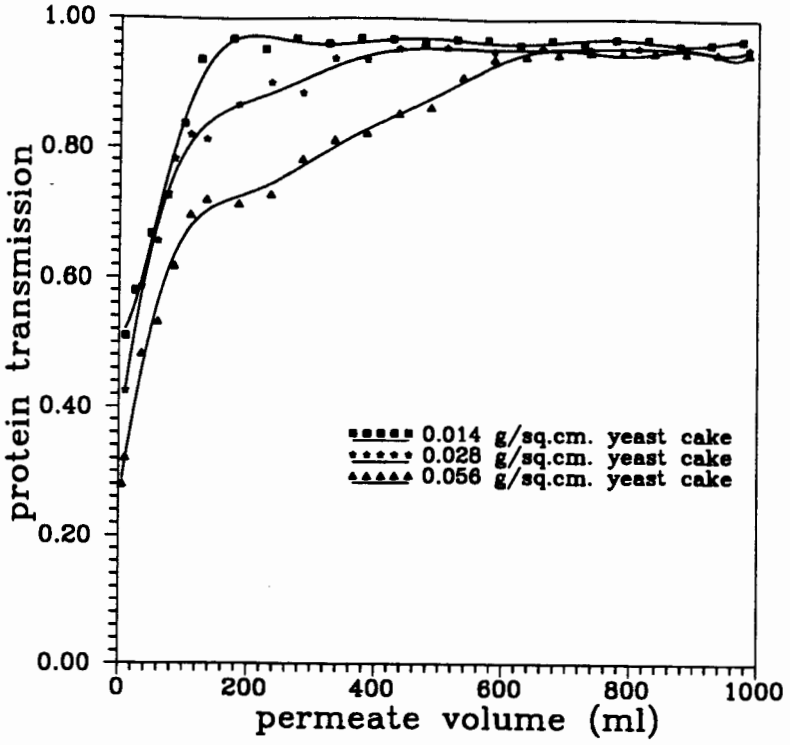


Figure 7: Protein transmission variation with the permeate volume collected, for 0.1 mg/ml BSA solution microfiltration in the presence of different thicknesses of vacuum filtered yeast cake layers.

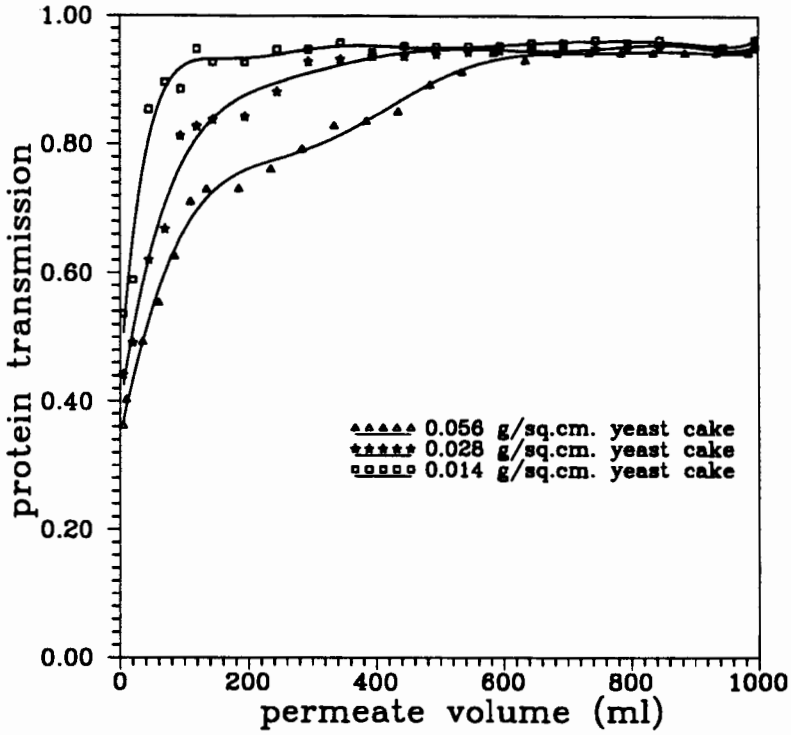


Figure 8: Protein transmission as a function of the permeate volume collected, for 0.1 mg/ml BSA solution microfiltration in the presence of different thicknesses of gravity preformed yeast cake layers.

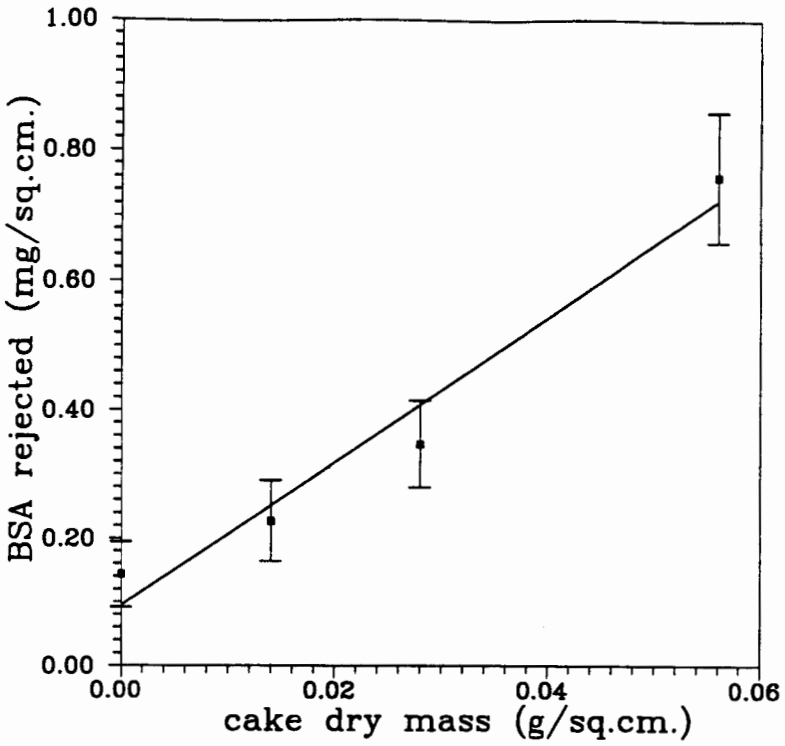


Figure 9: Plot of the mass of BSA rejected per unit external cake surface area (excluding BSA rejection due to the surface filtration) versus the cake dry mass per unit external surface area, for the vacuum filtered yeast cakes.

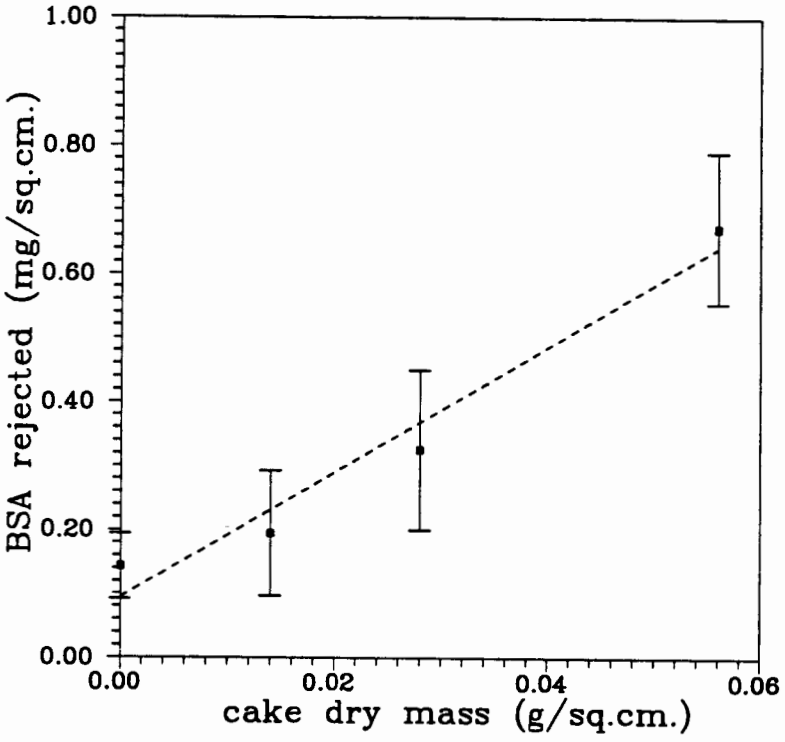


Figure 10: Mass of BSA rejected per unit external cake surface area (excluding BSA rejection due to the surface filtration) as a function of the cake dry mass per unit external surface area, for the yeast cake layers preformed by gravity settling.

MONITORING THE FATE OF TOLUENE AND PHENOL IN THE RHIZOSPHERE

N.Muralidharan[1], Lawrence C.Davis[2] and Larry E . Erickson[1]
Departments of Chemical Engineering[1] and Biochemistry[2]
Kansas State University
Manhattan,KS 66506

ABSTRACT

The beneficial effects of vegetation in the bioremediation of toluene and phenol were investigated individually in two pilot scale systems in which alfalfa was growing in the soil composed of sand and silt taken from the Kansas river valley near Manhattan, Kansas.

A closed tank having channels with dimensions of 10 cm width and 35 cm depth was filled with the aquifer soil materials mentioned above. Alfalfa plants have been growing in the aquifer soil materials for more than 9 months under fluorescent illumination with flowing subsurface water. The water source flowing into each of the channels was respectively saturated with toluene at 26 °C in one case and phenol at 500 ppm in the other. The subsurface flowing water simulated the conditions of the groundwater for the growing plants.

Concentrations of toluene and phenol in the subsurface water were monitored by 4 sampling wells spaced about 30 cm apart. The concentrations of toluene and phenol showed little change along the channel, but the concentrations were definitely lower than the inflow source solutions. However, most of the water and organic contaminants which entered the rhizosphere did not leave in the liquid effluent. Most of the water leaves by evapotranspiration while the toluene and phenol appear to be biodegraded.

Monitoring of the concentrations was done by measuring the UV absorbance of the contaminants in the solution. This was accomplished by extracting the contaminant into an organic phase to avoid the soil impurities in the drawn solutions from the sampling ports. The contaminant concentration in the solution phase was then determined by measuring the absorbance in the extracted phase at 261 nm and 263.5 nm for toluene and phenol respectively.

OBJECTIVES

1) To investigate the potential benefits of plants in the bioremediation of soil and groundwater contaminated with toluene and phenol separately.

2) To obtain experimental data for comparison with simulation models and to determine if the models realistically describe the organic contaminant fate.

3) To simulate experimentally the fate of toluene and phenol for a site contaminated with these hazardous organic compounds.

METHODS

Sample collection -

One ml of the contaminated solution was drawn from the monitoring wells using a syringe fitted with a long nylon tube. Four sampling wells along with input and output served as 6 points of observation.

Contaminant extraction -

The drawn solution was immediately extracted into an organic liquid phase[3]. Toluene was extracted into heptane and phenol into n-octanol. Toluene is a fairly volatile organic compound[2], hence its extraction had to be done quickly. For phenol 5 ml of buffer solution (pH=7) containing 0.1 molar phosphate was added before extraction.

Concentration measurements -

The concentrations of the contaminants were monitored using a Beckman Spectrophotometer model DU-2 with UV as the source of light. Toluene in heptane was measured at 261 nm wavelength[1,4] and phenol in n-octanol at 263.5 nm wavelength[1,4].

Water table depth measurements -

The depth of the water table in the tank at the sampling wells was measured using a pair of wires which were sent through the sampling well until it touched the surface of the water table. Upon contact with water an increased conductance was recorded by the ohm-meter.

Gas phase measurements -

Gas phase measurements were made using a FT-IR spectrometer and gas chromotography. Our thanks to C.Chaffin and V.P.Visser for assistance with the FT-IR measurements. Further results will be presented in the Hazardous Waste Research Conference 1993.

RESULTS

The water table level measurements showed that the levels were usually between 25-30 cm below the surface of the soil. In other words, it was from 5-10 cm high in the 35 cm deep tank. The water table was usually maintained by the amount of the inflow water solution.

In Tables 1 to 3 the concentrations of toluene and phenol are expressed in dimensionless form based on the inlet concentration. For toluene, the inflow concentration was the saturated concentration(515 mg/liter) at that temperature[5]. For phenol, the inflow concentration was 500 ppm by volume. Tables 1 and 2 are for toluene concentrations at slow and rapid flow rates respectively. Table 3 is for phenol.

TABLE 1

Dimensionless concentration of toluene in the saturated zone at five axial positions for **slow flow rates**. Inlet flow was water saturated with toluene. The input concentration was about 0.5 ml/liter(515 mg/liter).

Average daily flow rate (liters/day)		DIMENSIONLESS CONCENTRATION C^* C^* = concentration C / inlet concentration				
Inflow	Outflow	Port 1	Port 2	Port 3	Port 4	Port 5#
1.0	0.4	0.89	0.97	0.99	0.81	1.06
1.0	0.4	0.80	0.76	0.83	0.91	0.87
1.5	0.8	0.82	0.87	0.75	0.69	0.96
1.5	0.85	0.83	0.92	0.90	0.81	0.97

Each experimental sample was done at least 2 days apart.
Port 5 was connected to outflow container.

TABLE 2

Dimensionless concentration of toluene in the saturated zone at five axial positions for relatively **rapid flow rates**. The concentration was about 0.5 ml/liter(515 mg/liter). Values were obtained during recharge of a depleted chamber. (water level was below 33 cm on day previous to expt. 1)

Experiment No:+	Average daily flow rate (liters/day)		DIMENSIONLESS CONCENTRATION C* C* = concentration C / inlet concentration				
	Inflow	Outflow	Port 1	Port 2	Port 3	Port 4	Port 5#
1A	3.0	1.5	0.94	0.96	0.94	1.06	1.02
1B	3.0	1.5	1.06	0.91	0.86	1.07	0.96
2A	3.5	1.1	0.91	0.92	1.03	0.79	0.99
2B	3.5	1.1	0.94	0.87	0.98	0.80	0.95

+ Experiments A and B were repeated samples taken 40 mins apart on the same day.

Port 5 was connected to outflow container.

TABLE 3

Dimensionless concentration of phenol at five axial positions. Inlet flow was water with phenol. The concentration was 0.5 ml/liter.

Average daily flow rate (liters/day)		DIMENSIONLESS CONCENTRATION C* C* = concentration C / inlet concentration				
Inflow	Outflow	Port 1	Port 2	Port 3	Port 4	Port 5#
1.0	0.1	0.89	0.91	0.66	0.98	0.02
0.6	0.0	0.77	0.76	0.24	-	-
0.6	0.0	0.85	0.79	0.25	-	-
1.4	0.0	0.96	0.89	0.82	0.77	0.02+
1.2	0.3	0.99	0.89	0.84	0.77	0.03

+ No measurable effluent flow. - No sample was collected.

Port 5 was connected to outflow container.

Using the inflow and outflow rates and concentrations a steady state mass balance was made across the tank to estimate the levels of loss of contaminants and water per day.

Tables 4 and 5 show the fractions of contaminants and water which do not leave in the effluent flow for toluene. Table 6 contains similar results for phenol.

TABLE 4

Fraction of water and toluene which do not leave in the effluent for **slow flow rates**.

Average daily flow rates (liters/day)		C^*	fraction of toluene lost	fraction of water lost
Inflow	Outflow	Port 5		
1.0	0.40	1.06	0.58	0.6
1.0	0.40	0.87	0.65	0.6
1.5	0.80	0.96	0.49	0.47
1.5	0.85	0.97	0.45	0.43

TABLE 5

Fraction of water and toluene which do not leave in the effluent for **rapid flow rates**.

Average daily flow rate (liters/day)		C^*	fraction of toluene lost	fraction of water lost
Inflow	Outflow	Port 5		
3.0	1.5	1.02	0.49	0.50
3.0	1.5	0.96	0.52	0.50
3.5	1.1	0.99	0.69	0.69
3.5	1.1	0.95	0.70	0.69

TABLE 6

Fraction of water and phenol which do not leave in the effluent.

Average flow rate (liters/day)		C^*	fraction of phenol lost	fraction of water lost
Inflow	Outflow	Port 5		
1.0	0.1	0.02	0.92	0.87
0.6	0.0	-	1.00	1.00
0.6	0.0	-	1.00	1.00
1.4	0.0	0.02	1.00	1.00
1.2	0.3	0.03	0.97	0.85

FT-IR spectrometer measurements showed undetectable levels of toluene in the gas phase during several attempts to monitor accumulation in the closed chamber. This indicated that most of the toluene which partitions into the gas phase in the soil is biodegraded before it diffuses from the surface of the soil. Phenol was not sufficiently volatile to make such measurements reliable.

CO_2 diffusion from the soil into trapping containers indicated significant rates of biomass respiration occurring in the soil. Rates were higher near the inlet of the contaminant than near the outlet.

DISCUSSION

The dimensionless concentration shown in Tables 1, 2 and 3 show a slight decrease in the level of the concentration as compared to the input concentration in most of the positions for each of the contaminants. Further, the plants remove a significant fraction of the water through evapotranspiration as indicated in Tables 4 to 6. This indicates that a significant amount of the contaminants is lost with water along the channel.

Additionally, for toluene, the levels of the concentration varied only slightly along the channel. In other words, spatial and temporal variability were low in each case for toluene as compared to phenol. This meant that the loss of toluene was uniform and similar in amount, to the loss of water, by evapotranspiration, along the length of the channel. The amount of loss of toluene is shown in Tables 4 and 5. However, for phenol, the levels of concentration varied substantially along the channel. This meant that a significant part of phenol is lost. The fraction of contaminant which leaves in the effluent depends on the quantity of the solution which flows out of the system.

The vapor pressure of toluene and phenol[2] are about 30 mm and 1 mm of Hg respectively at 26°C (room temperature). The vapor pressure of water at room temperature is about 30 mm of Hg. Toluene is a more volatile organic component as compared to phenol. Toluene has a similar tendency as water to vaporize into the gas phase. This is also evident from Tables 4 and 5 which show almost similar fractions of loss of toluene and water.

In order to study the aerobic degradation, a variation of the saturated zone in the tank, and hence the unsaturated zone, was undertaken. The ability to change the inflow rate facilitated this study. The variation in the inflow rate was achieved by changing the level in the inlet storage container. A decrease in the elevation at the source caused a decrease in the level of the water table in the tank. Consequently, the vadose zone of the aquifer increased. This increase in unsaturated zone possibly resulted in more atmospheric oxygen diffusion into the soil. The increased availability of O_2 probably resulted in enhanced bioremediation. This is evident from the lower values of C^* in Table 3 as compared to the values in Table 2.

The higher values of C^* for toluene for rapid flow rates could be attributed to the fact that in the saturated zone the biodegradation of the contaminants is limited by the availability of oxygen. Degradation of toluene is presumbly an aerobic process initiated by microbial mixed function oxidases. Phenol can be degraded under anaerobic conditions.

CONCLUSIONS

It is evident from the results that a significant fraction of the inflow contaminant was not present in the effluent. Even when there was no outflow of effluent the concentration of contaminant in the aqueous phase did not increase. This indicates that the plants were not selectively excluding the contaminant from the transpiration stream over the long term, or that any such exclusion was being countered by increased degradation in the rhizosphere - vadose zone. Moreover, the decrease in the level of the contaminants from the saturated zone to the atmosphere above the soil supported the observation that most of the contaminant is lost in the plant or in the vadose zone of the soil. This loss appears to be due to biodegradation of the contaminants.

ACKNOWLEDGEMENTS

The work presented in this was partially supported by the US EPA under assistance agreements R-815709 and R-919653 to the Great Plains-Rocky Mountain Hazardous Substance Research Center for regions 7 and 8 under project 90-13. It has not been submitted to the EPA for peer review and therefore may not necessarily reflect the views of the agency and no official endorsement should be inferred. The U.S. Dept. of Energy, Office of Environmental Restoration and Waste Management, Office of Technology Development and the Center for Hazardous Substance Research also provided partial funding.

REFERENCES

1. Atlas of Spectral Data and Physical Constants for Organic Compounds. Vol. 4. Edition 2. Editors-J.G.Grasselli and W.M.Ritchey. 74 and 642. CRC Press; Boca Raton, FL, 1975.
2. Chemical Engineers' Handbook. Edition 5. Editors-R.H.Perry and C.H.Chilton. 3-40, 3-43, 3-58, 3-60. The McGraw-Hill, NY, 1973.
3. Groundwater Chemicals Desk Reference, J.H.Montgomery and L.M.Welkom. 473-477 and 501-506. Lewis Pub., Boca Raton, FL, 1990.
4. Handbook of Spectroscopy. Vol 1 and 2. Editor-J.W.Robinson. 194 and 208. CRC; Boca Raton, FL, Press 1974.
5. Solubilities of Inorganic and Organic compounds. Vol. 1 and 2. Editors-H.Stephen and T.Stephen. 62, 67. The MacMillan Comp., NY, 1963.

Hydrodynamic Motions Associated with Bubble Coalescence and Breakup

T. Y. Yiin, L. A. Glasgow, and L. E. Erickson
Department of Chemical Engineering
Kansas State University
Manhattan, KS 66502-5102

Abstract

Mechanisms of cellular damage must be identified in order to facilitate rational design and scale-up of reactors for production of animal glycoproteins such as interferon. The loss of cellular viability in both insect and animal cell cultures has been attributed to interfacially driven fluid motions. The important events may include bubble formation and detachment at the sparger, bubble coalescence and breakage, and bubble disengagement at the free surface, entailing film thinning, film rupture, and film droplet ejection.

Experimental observations have been made using high-speed macrovideography to investigate the magnitude of hydrodynamic forces associated with bubble-induced fluid motions. Liquid media employed to assess the effects of viscosity included distilled water and 48 and 72 weight percent solutions of glycerol (μ = 5 and 17 cp, respectively). In the 48% glycerol solution, for example, once bubble coalescence began, the average velocity of the expanding rim was around 52 cm/sec, and the time required for coalescence was about 0.003 - 0.008 sec. The time required for bubble breakup at the surface was about 0.001 - 0.007 sec; the rim expanded very quickly for the first ~0.001 sec at about 3 to 5 m/sec, subsequently the velocity rapidly diminished to around 0.2 to 0.6 m/sec. The measurement of these velocities indicated that some bubble-induced fluid motions may contribute to the loss of cellular viability through generation of very large shear stresses.

Introduction

Airlift reactors have a number of advantages over more traditional stirred tank reactors; these features include effective mixing at lower average shear stress, less contamination, sufficient oxygen transfer, and good temperature control. Accordingly, airlift reactors have found a wide variety of culture applications, such as the production of monoclonal antibodies with hybridoma cells. However, some instances of cellular damage have been reported; speculation on the causes of loss of cellular viability has centered on interfacially driven fluid motions. The important events may include bubble formation and detachment at the sparger, bubble coalescence and breakage, and bubble disengagement at the free surface, entailing film thinning, film rupture, and film droplet ejection. However, there is no consensus regarding the effects of those above-mentioned phenomena upon cellular viability in cultures with direct sparging.

Tramper et al. (1986) reported that insect cells could not grow in an airlift reactor; they attributed the loss of cellular viability to bubbles rising through the liquid and bubbles bursting at the surface of the suspension. Handa et al. (1987) studied the culture of hybridoma and myeloma cells in small bubble columns with sintered aeration disks; they found cell growth suffered in response to small bubbles coupled with increased gas rates. Handa-Corrigan et al. (1989) presented their experimental data to support the hypothesis that cell death is associated only with

the region of bubble disengagement; they studied cell death in sparged systems for suspended mammalian cells cultured in bubble column fermenters. They proposed two mechanisms of cell damage: one due to rapid oscillation caused by bursting bubbles, and the other due to shear caused by draining liquid films in foams. Kunas and Papoutsakis (1990) determined apparent growth rates for hybridoma cultures in agitated bioreactors with and without bubble entrainment; they found that cell damage is associated with vortex formation accompanied by bubble entrainment and breakup when a gas phase is present.

When bubble breakup occurs in a low-viscosity liquid, the film droplets are produced very rapidly, whereas jet drops are formed at the end of bubble film retraction by the pencil-like jet which rises from the bottom of the bubble cavity. Afeti (1987) investigated bubble breakup at low-viscosity liquid surfaces by using a laser-based inline holographic technique; the major feature included a very powerful and coherent source of high intensity light so that the spatial location of the drops with respect to each other as well as their true sizes and shapes could be directly and accurately determined.

In the present study of bubble-induced fluid motions as sources of cellular damage, we have investigated bubble phenomena at the disengagement surface by using Kodak Ektapro high-speed macrovideography; the data collected so far include the expanding rim velocity during both bubble breakup and coalescence, and the velocities of ejected film droplets.

Experimental Apparatus

A plastic container with the dimensions of 12x19x13 cm was used for study of bubble coalescence and disengagement; air was delivered by pump via a tube with inner diameter of 3.57 mm which was immersed at the bottom of the container, bubbles would rise to the surface, and either coalesce or breakup. The Kodak High-Speed Imager, positioned in front of the container, recorded the bubble events from above at a speed of 1,000 frames per second; the lighting system was provided by high-power flood lights, 20 cm away from both the right and left sides of the container. Some preliminary video observations of bubble phenomena were also performed in a 3-liter, acrylic plastic, split-column airlift reactor where air was introduced through a sieve plate sparger with holes of 1 mm diameter. To assess the effects of viscosity and surface tension, the liquid media used included distilled water, 48 and 72 weight percent solutions of glycerol (μ = 5 and 17 cp, respectively), and a surfactant solution.

Results and Discussion

An example series of photographs of the breakage of a single 19.8 mm diameter bubble in the air-water system is provided as Figure 1 (1/1,000 s frame rate); the bubble film rupture initiation was observed at frame 15, which shows a hole diameter of 5.22 mm. The hole expanded very quickly with the rim velocity increasing to 397 cm/sec within the first 0.002 sec, and subsequently diminishing to 135 cm/sec as shown in Figure 2. The observed expanding rim velocities for various size bubbles are included in Figure 2; the maximum velocity always occurred within 0.002 sec, subsequently the velocity diminished to about 200 cm/sec at 0.005 sec after rupture. Figure 3 contains similar data for expanding rim velocity in the 48% glycerol solution; the velocity changed with the same tendency as in distilled water, quickly reaching

the maximum value, and subsequently diminishing over a period of 0.004 or 0.005s. However, the maximum rim velocity was somewhat diminished. At glycerol concentrations of 72%, we observed that the time required for breakup events decreased to less than about 0.003 sec; clear observation could not be made with the high-speed video system.

Following bubble film rupture initiation, film droplet ejection was observed accompanying hole expansion; the droplet velocities ranged from 100 cm/sec to 500 cm/sec. Figure 4 contains the cumulative fraction of film droplet ejection velocities from the surface of distilled water; the most likely values of the droplet velocities are between 200 cm/sec and 350 cm/sec. In contrast, Figure 5 contains the cumulative fraction of film droplet ejection velocities from surface of a surfactant solution; the range of droplet ejection velocities is between 150 cm/sec and 600 cm/sec, and more than 50% of the detected values are below 300 cm/sec.

Afeti (1987) described the mechanism of bubble breakup from the analysis of scores of holograms: first film thinning occurred due to the combined effect of gravity and surface curvature, leading to rapid fluctuations in the film thickness, and finally the thin areas served as initial rupture points. Afeti reported the time between the initial perforations in the film cap and the production of the first film drops was about 150 μs; film drop formation was completed in the next 150 μs with the disintegration of all the filaments into droplets. However, we observed film droplets from rupture of large bubbles produced after a couple of ms and the droplet production process lasted for about 2 ms. Afeti also observed absence of jet drop formation for bubbles larger than about 6 mm in diameter; our video observations did not show any jet drops, but of course our typical bubble diameters were on the order of 15-40 mm. Bubble coalescence events were observed in the 3-liter airlift reactor as the video system was easily set up to observe these events. For 48% glycerol solutions, two bubbles approached each other, the intervening film would drain for a couple of ms prior to coalescence. The illustration of rim velocity vs time is shown in Figure 6. Actually, bubble breakup and bubble coalescence events are similar in that the hole initiating the process expands very quickly.

Conclusions

These experimental results clearly show the importance of bubble coalescence, bubble breakup, and film droplet ejection with respect to generation of large, local velocities. The high-speed video system produced qualitative and quantitative information in nearly all of the recorded cases. The results have led us to further consider the speculation as to which interfacial phenomena are especially harmful to cellular entities. Chalmers et al. (1991) has investigated the interactions of insect cells with rupturing bubbles by a microscopic, high-speed video system; they observed the adsorption of cells onto the bubble film. Accordingly, possible cellular enrichment at the gas-liquid interface may have an amplifying effect upon the importance of interfacially driven fluid motions at the surface of sparged cultures. Based on the data obtained in this study, however, we believe the bubble breakup and bubble coalescence processes at the sparger are energetic enough to warrant additional attention.

Acknowledgement

This research was partially supported by the National Science Foundation and the State of Kansas through the K*STAR Experimental Program to Stimulate Competitive Research.

References

1. Tramper, J., Williams, J.B., Joustra, D., and Vlak, J.M. 1986. Shear sensitivity of insect cells in suspension. Enzyme Microb. Technol. 8:33-36.

2. Handa, A., Emery, A.N., and Spier, R.E. 1987. On the evaluation of gas-liquid interfacial effects on hybridoma viability in bubble column bioreactors. Develop. Biol. Stan. 66:241-253.

3. Handa-Corrigan, A., Emery, A.N., and Spier, R.E. 1989. Effect of gas-liquid interfaces on the growth of suspended mammalian cells:mechanisms of cell damage by bubbles. Enzyme Microb. Technol. 11:230-235.

4. Kunas, K., and Papoutsakis, E. 1990. Damage mechanisms of suspended animal cells in agitated bioreactors with and without bubble entrainment. Biotechnol. Bioeng. 36:476-483.

5. Afeti, G.M. Bubble break-up at a low-viscosity liquid surface. Cavitation and Multiphase Flow Forum:26-29(1987); presented at 1987 ASME Applied Mechanics, Bioengineering, and Fluids Engineering Conference (Cincinnati).

6. Chalmers, J., and Bavarian, F. 1991. Microscopic visualization of insect cell-bubble interactions II:The bubble film and bubble rupture. Biotechnol. Prog. 7:151-158.

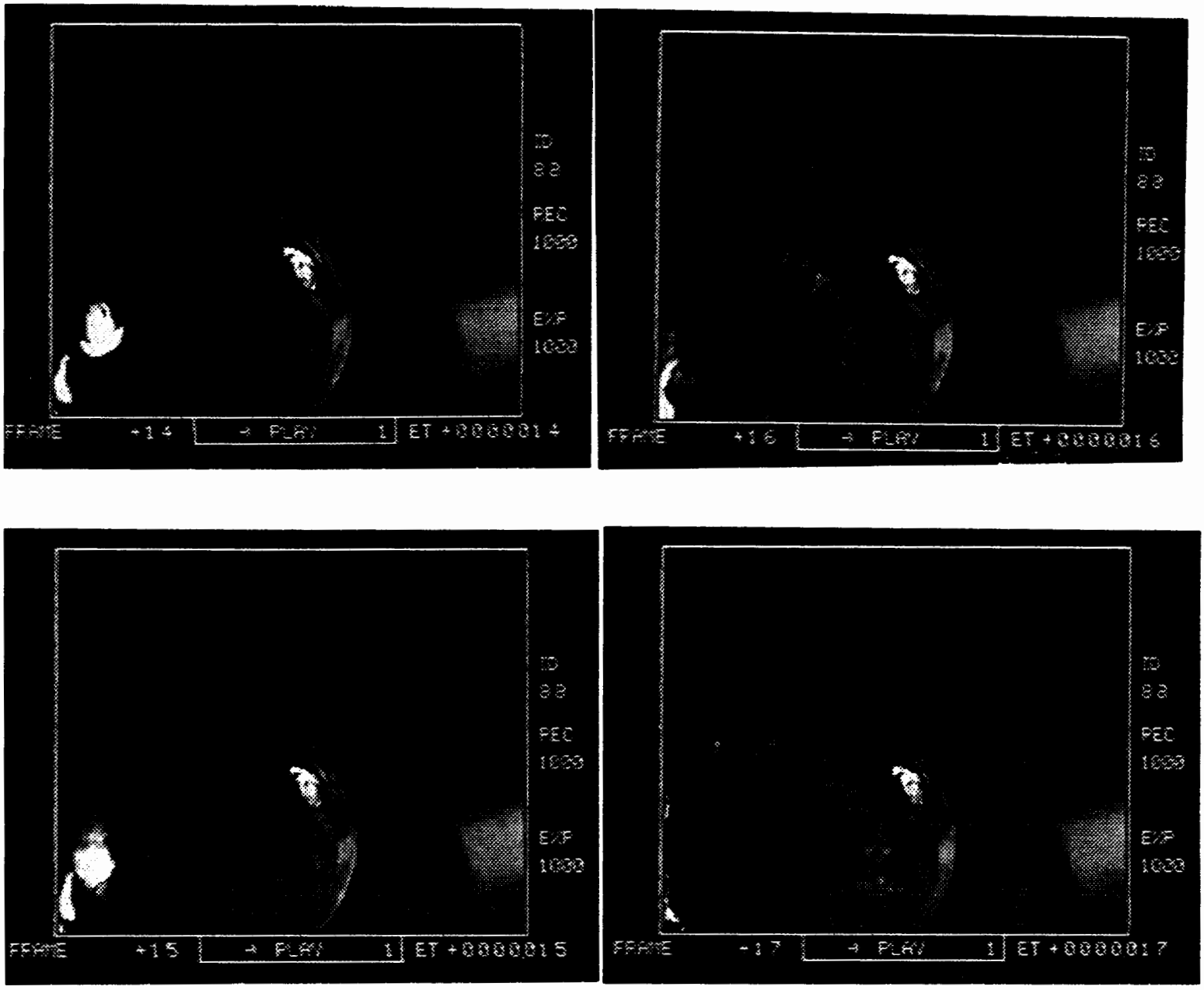


Figure 1A: A series of photographs of the breakage of a single 19.8 mm diameter bubble in the air-water system; frame 14 to 17.

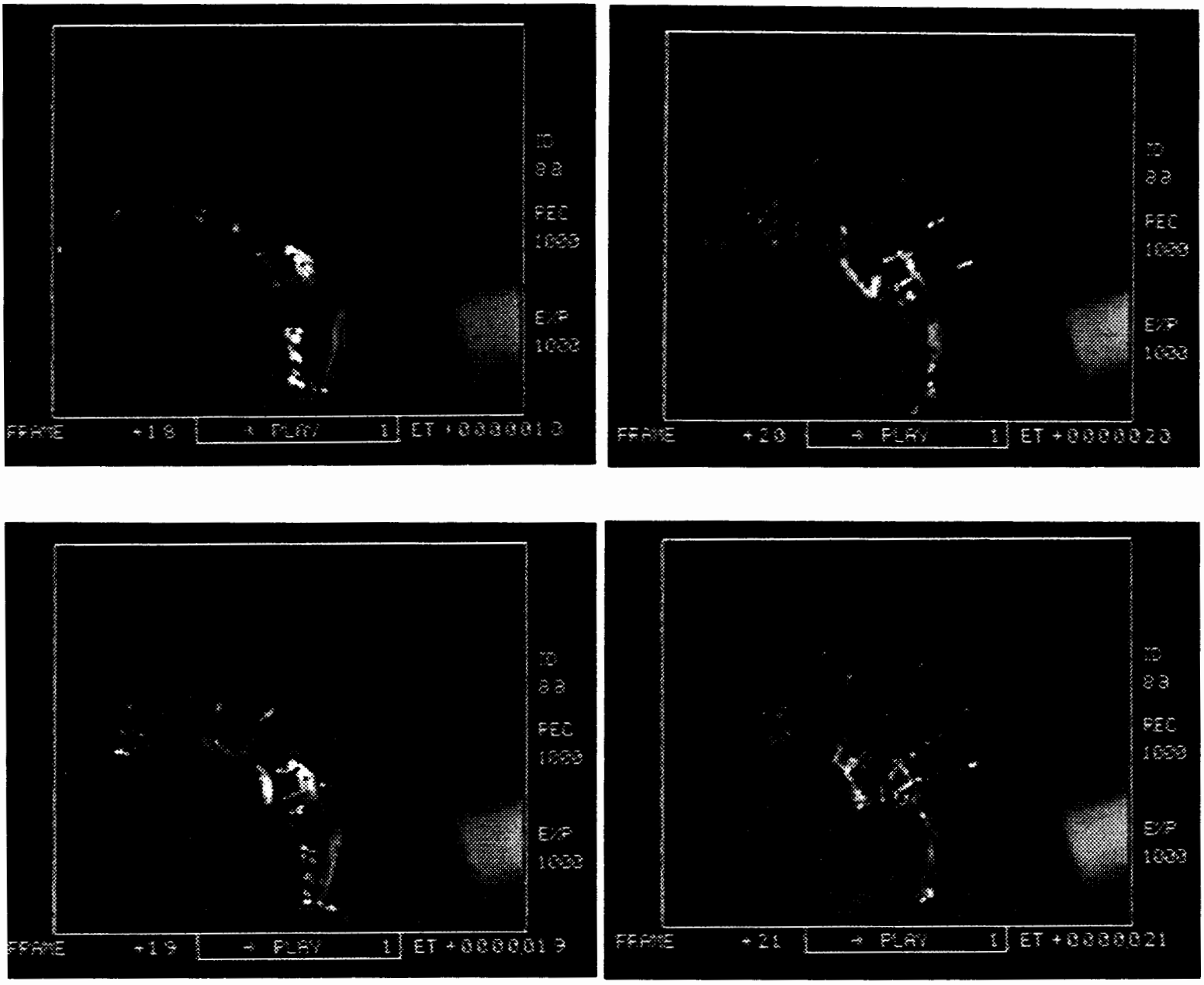


Figure 1B: A series of photographs of the breakage of a single 19.8 mm diameter bubble in the air-water system; frame 18 to 21.

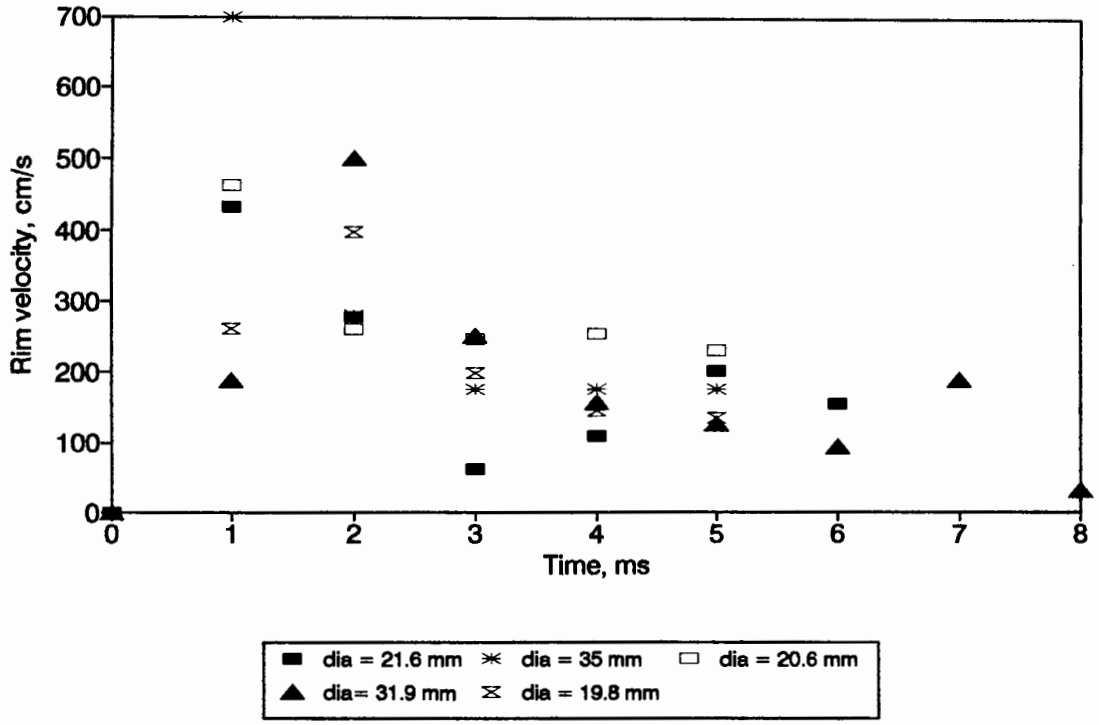


Figure 2: Estimated expanding rim velocities for breakup of bubbles with different diameters in an air-distilled water system

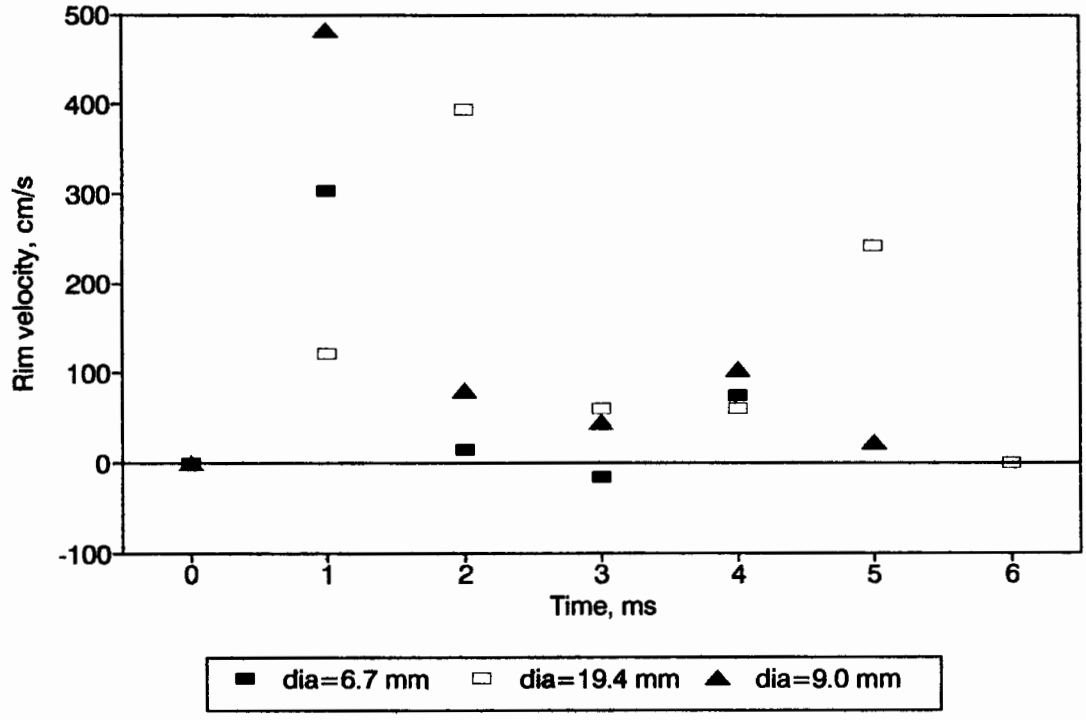


Figure 3: Estimated expanding rim velocities for breakup of bubbles with different diameters in 48% glycerol solution

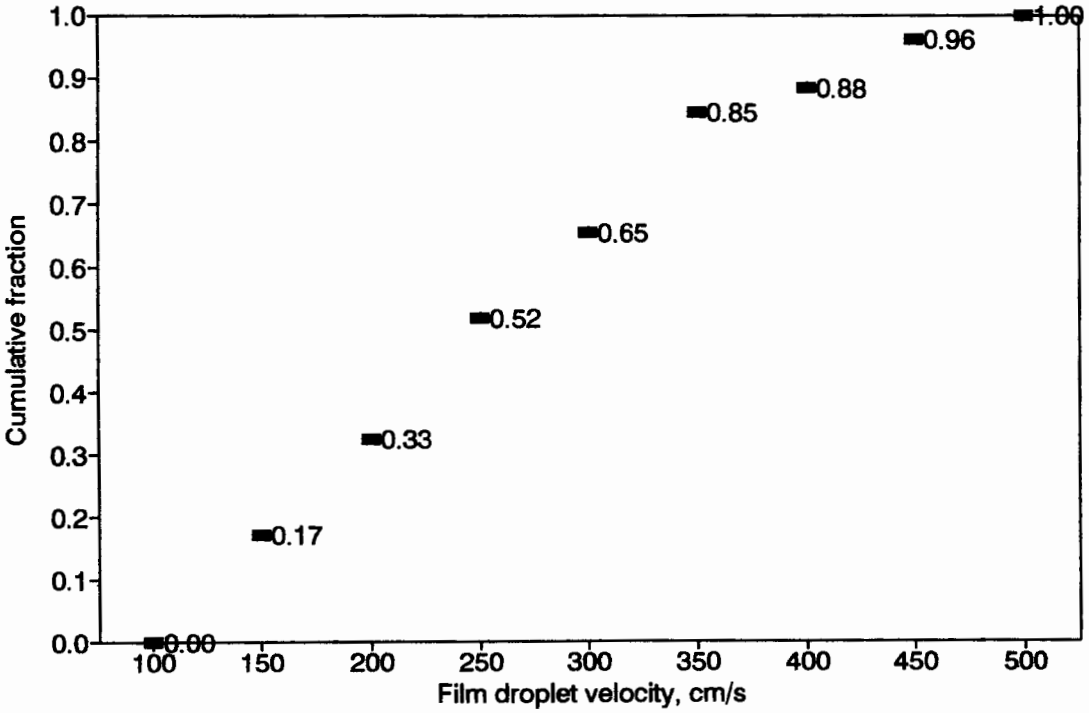


Figure 4: Film droplet ejection from surface of distilled water

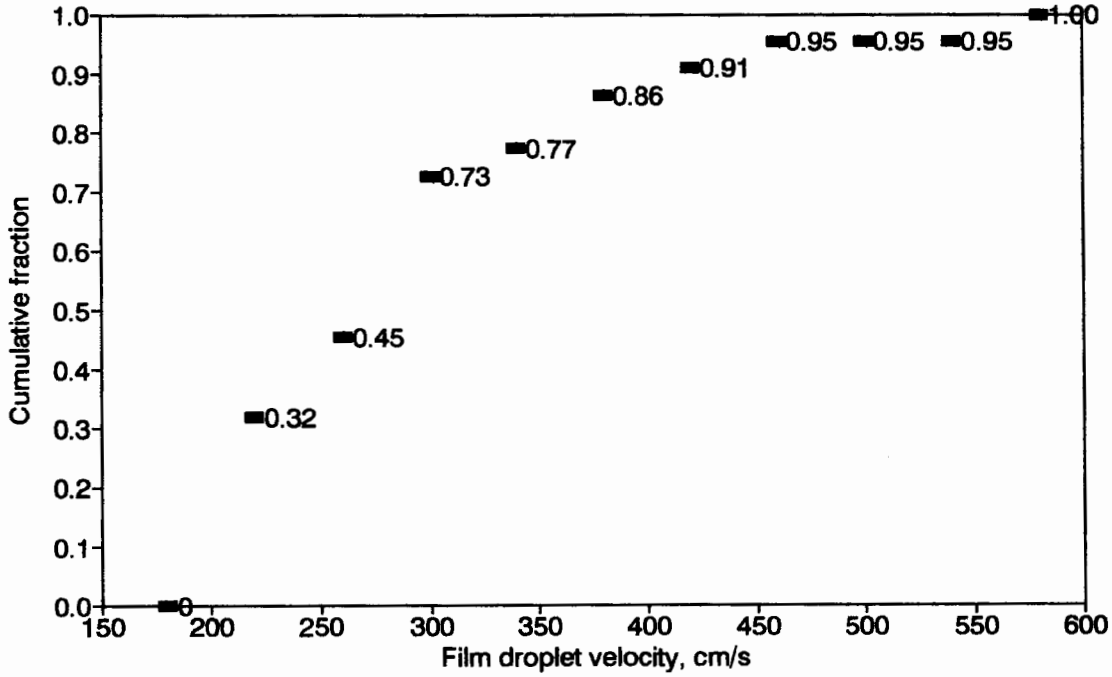


Figure 5: Film droplet ejection from surface of surfactant solution

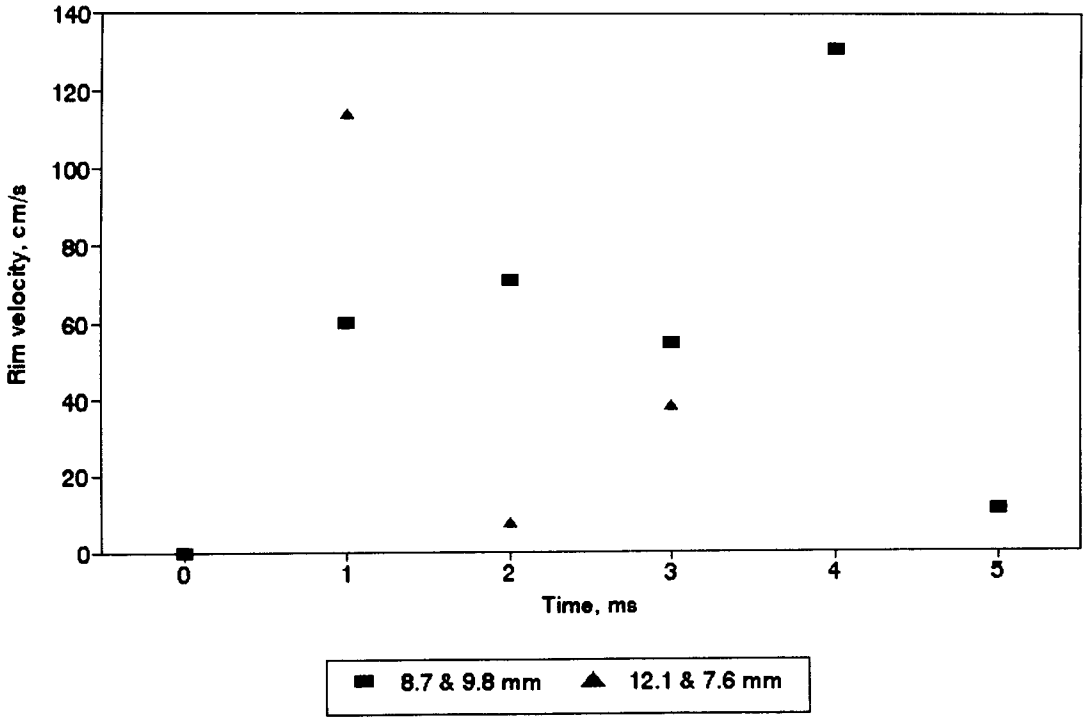


Figure 6: Rim velocity when two bubbles coalesce in 48% glycerol solution

EXPRESSION AND PURIFICATION OF α-HUMAN ATRIAL NATRIURETIC PEPTIDE IN *ESCHERICHIA COLI* BY FUSION WITH L-ASPARAGINASE

Nien-Tung Ma and Roger G. Harrison
School of Chemical Engineering and Materials Science,
University of Oklahoma, Norman, Oklahoma 73019

ABSTRACT

Expression vectors have been constructed in which the gene for α-human atrial natriuretic peptide is joined to the gene for L-asparaginase from *Erwinia chrysanthemi* with a factor Xa protease linker gene. L-asparaginase, with an isoelectric point 8.6, was selected as a partner in this fusion protein because of the possibility of developing a simple and inexpensive affinity purification system using immobilized L-asparagine. Construction with the native L-asparaginase signal correctly targets the fusion protein to the periplasm in *Escherichia coli*, while construction without the signal sequence causes expression only in the cytoplasm of *E. coli*. This fusion protein in crude extract can be purified by either affinity chromatography or cation exchange chromatography. This fusion system can be potentially used for the expression and purification of many other eucaryotic peptides in *E. coli*.

INTRODUCTION

For the past few years, eucaryotic proteins and peptides have been successfully expressed and purified in procaryotic cells by the gene fusion method. This procedure is employed to produce these valuable proteins and peptides, which are useful in pharmaceutical, agricultural, and industrial applications but limited by their natural availability. Some systems have been developed for production fusion proteins in *Escherichia coli*. In general, each system involves the connection of a favorite gene to a carrier gene with a protease recognized linker sequence, so that a single step affinity purification is facilitated (reviewed by Marston, 1986 and Uhlén, *et al.*, 1990). Some examples of these systems are the following: Pre-S2 peptide of hepatitis B envelope protein has been expressed in a fusion protein with β-galactosidase and purified by *p*-aminophenyl-β-D-thiogalactoside Sepharose to a level of 90% homogeneity (Offensperger, *et al.*, 1985). An insulin-like growth factor has been fused to IgG-binding domains of protein A and purified on an IgG-Sepharose column (Löwenadler, *et al.*, 1987). An alkaline phosphatase signal sequence has been used to synthesize human epidermal growth factor in the periplasmic fraction of *E. coli* cells (Oka, *et al.*, 1985). A maltose binding protein fusion (Maina., *et al.*, 1988) and a glutathione S-transferase fusion system (Smith, *et al.*, 1988) have been constructed to produce proteins that have been affinity purified with cross-linked amylose and glutathione agarose, respectively. Disadvantages include the requirement for denaturation and refolding in some systems, the costs of ligands, and the efficiency of the yield.

The present study was initiated to develop a new method which utilizes recombinant DNA technology and affinity separation to produce peptides in a cost effective fashion. L-asparaginase from *Erwinia chrysanthemi* (NCPPB1066) is a tetramer consisting of identical subunits of molecular weight 35,000. This *Erwinia* L-asparaginase gene has been cloned and expressed (up to 6% of soluble protein) in *E. coli* (Gilbert *et al.*, 1986) and sequenced

(Minton *et al.*, 1986). One attractive feature of *Erwinia* L-asparaginase is an unusually high isoelectric point (pI) of 8.6. This high pI offers the possibility of a high degree of purification by ion exchange chromatography. A large scale process was developed for the purification of L-asparaginase on an L-asparagine Sepharose 6 Fast Flow affinity column (Lee *et al.*, 1989). In the present study, recombinant DNA technology was used to join eucaryotic α-human atrial natriuretic peptide (de Bold, 1985) to *Erwinia* L-asparaginase through a cleavable linker sequence recognized by blood coagulation factor Xa to form a fusion protein. Fusion proteins with or without the L-asparaginase signal sequence were cloned into a pKK223-3 vector and expressed under the control of the IPTG (isopropyl-β-D-thiogalactopyroanoside) inducible *tac* promoter in *E. coli*. The fusion protein synthesized in the cytoplasm was purified either by affinity chromatography or by cation exchange chromatography. The purified fusion protein was analyzed by SDS-polyacrylamide gel electrophoresis and immunoblotting.

MATERIALS AND METHODS

Bacterial strains and plasmids *E. coli* strain JM101 (*supE*, *thi*, Δ(*lac-proAB*), [F', *traD*36, *proAB*$^+$, *lacI*^{q}, *lacZ*ΔM15]) was used as a host cell for the maintenance and propagation of M13 DNA and its recombinants. *E. coli* SDM (*hsdR*17, *mcrAB*, *recA*1, *supE*44, *Tet*^{r}, Δ(*lac-proAB*), [F' *traD*36, *proAB*$^+$, *lacI*^{q}, *lacZ*ΔM15]), was used as a host cell for *in vitro* mutagenesis and purchased from USB. JM105 (*thi*, *rspL*, *endA*, *sbcB*15, *supE*, *hdsR*4, Δ(*lac-proAB*), [F', *traD*36, *proAB*$^+$, *lacI*^{q}, *lacZ*ΔM15]) was used for the expression of fusion proteins and obtained from Pharmacia. JM83 (F$^-$, *ara*, Δ(*lac-proAB*), *rpsL*, Φ80*lacZ*ΔM15]) harboring plasmid pASN326 was from Microbial Technology Laboratory, PHLS, Centre for Applied Microbiology and Research in England. Plasmid pASN326 containing *Erwinia chrysanthemi* NCPPB 1066 L-asparaginase gene has an ampicillin-resistance gene. Plasmid pKK223-3 with the *tac* promoter was used as the expression vector and was obtained from Pharmacia. The α-hANP was synthesized, cloned into M13mp18 vector, and sequenced by Biosynthesis, Inc. (Lewisville, Texas).

DNA sequencing The dideoxynucleotide DNA sequencing was performed by *Bacillus stearothermophilus* (*Bst*) DNA polymerase in a radiolabled two-step sequencing reaction (Chissoe *et al.*, 1991). Each reaction contained 500-750 ng of single-stranded DNA, 2 μl of *Bst* reaction buffer [500 mM Tris-HCl (pH 8.5), 150 mM $MgCl_2$], 2 μl of nucleotide extension mix (15 μM dCTP, c^{7}dGTP, and dTPP), either 1 μl of M13 (-40) universal 17-mer sequencing primer (2.5 ng/μl) or 1-2 μl of a 20 ng/μl 9-mer primer (Table 1), 0.5-1.0 μl α[^{32}P]ATP (10 μCi/μl) (Amersham), and 1 μl *Bst* DNA polymerase (0.1 U/μl) (Bio-Rad) in a final volume of 12 μl. After the reactions were extended for 2 min at 67°C and briefly centrifuged, 2.5 μl aliquots were removed and added to 2.5 μl of termination mix (8 μM limiting dNTP, 164 μM nonlimiting dNTP, and 660 μM ddATP, 400 μM ddCTP, 540 μM ddGTP, or 600 μM ddTTP). Following an incubation at 67°C for 10 min and a brief centrifugation, the reactions were stopped by the addition of 4 μl of dye/formamide/EDTA solution (0.3% xylene cyanol FF, 0.3% bromophenol blue, 10 mM EDTA in deionized formamide) and incubated at 100°C for 5-7 min. The reaction mix then was loaded into a polyacrylamide gel and electrophoresized at 22 mA. After the front dye went to the bottom, the power was disconnected.

Site-directed *in vitro* mutagenesis Oligonucleotide-directed *in vitro* mutagenesis (Vandeyar

et al., 1988) was performed to create the double-site mutation. The annealing reaction was carried out by mixing phosphorylated mutagenic primers (20 pmole each, Table 1) with 4 µg of single-stranded DNA template to 10 µl, heating at 65°C for 5 min, then cooling slowly at room temperature. To this mixture, 2 µl of synthesis buffer [10 mM Tris-HCl (pH 7.5), 20 mM 1,4-dithiothreitol (DTT), 5 mM dATP, 5 mM dGTP, 5 mM dTTP, 5 mM 5-methyl-dCTP, and 10 mM ATP], 1µl of T7 DNA polymerase (2.5 units/µl), 6 µl of T4 DNA ligase (1 unit/µl), and 1 µl of water were added and incubated at room temperature overnight to extend the primers in the presence of 5-methyl-dCTP. The nonmethylated parental strand was nicked by the incubation with *Sau*3A and *Hha*I. This digestion would leave the new methylated strand unattacked and only nick the nonmethylated parental strand, which was subsequently removed by exonuclease III. This exonuclease III digested mixture was transfected with calcium chloride treated *E. coli* SDM cells. The single-stranded DNA was isolated to screen mutants by using DNA sequencing method.

Polymerase chain reaction PCR was performed in a final volume of 100 µl containing 50 ng of double-stranded vector DNA, 6 mM of each primer (Table 1, PEWS/PPST for the construction of pCP32 and PENL/PPST for the construction of pCA919), 0.2 mM each deoxynucleoside triphosphate, reaction buffer [10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$] and 2 units *Taq* DNA polymerase (AmpliTaq, Perkin-Elmer). One drop of mineral oil was layered over the reaction mixture. The PCR was run on a programmable thermal cycler (Perkin-Elmer Cetus). Conditions were as follows: 4 min at 95°C, followed by 30 cycles consisting of 1 min denaturation step at 94°C, 1 min annealing step at 55°C, 2 min polymerization at 72°C, and then 10 min held at 72°C. DNA from PCR amplification was purified using GENECLEAN II kit (BIO 101 Inc.).

Expression of the fusion genes and cell fractionation JM105 cells harboring the vectors of fusion genes were inoculated into 5 ml LB/ampicillin (50 µg/ml) medium (10 ml test tube) and incubated at 37°C overnight. A 1:100 dilution of the overnight culture was inoculated into 50 ml LB/ampicillin medium (250 ml flask) and grown with shaking at 37°C. At an optical density of 0.4 at 600 nm, IPTG was added to a final concentration of 1 mM. The growth temperature was reduced to 30°C, and growth was continued for 3 h with shaking. The cells were harvested by centrifugation at 5000X g for 15 min at 4°C and resuspended in 5 ml of ice cold lysis buffer [20% sucrose, 0.2 mM DTT, 30 mM Tris-HCl (pH 8.0), and 1 mM EDTA]. They were then subjected to osmotic shock (Neu *et al.*, 1965). After periplasmic proteins were isolated by centrifugation, the pellets were resuspended in 2 ml of 10 mM Tris-HCl (pH 7.5) and 0.2 mM DTT and sonicated at 60W. DNAse and RNAse were added to the resuspended protein pellets and incubated at room temperature for 30 min. Cytoplasmic proteins were present in the supernatant, and membrane proteins and inclusion bodies were present in the pellet.

Asparaginase activity assay In this study, asparaginase activity was determined by direct Nesslerization as described previously (Wriston, 1970) with modification. Crude extract in sodium borate buffer (0.05 M, pH 8.5) containing 0.005% bovine serum albumin was added to borate buffer to a volume of 1.5 ml. A 0.5 ml volume of 0.04 M L-asparagine substrate was added to initiate the reaction with incubation of the reaction mixture at 37°C for 30 min. The reaction was stopped by adding 0.5 ml of 15% trichloroacetic acid and spun to remove precipitates. The supernatant was transferred quantitatively to water to a volume of 4.5 ml and reacted with 0.5 ml Nessler's reagent. After a further 15 min incubation at room temperature, the mixture was spectrophotometrically measured at 480 nm using a standard

curve prepared with ammonium sulfate. An enzyme unit was defined as the enzyme activity that forms 1 μmol of ammonia per minute at 37°C.

SDS-PAGE AND IMMUNOBLOTTING Proteins were examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmili, 1970). The gels were stained with Coomassie blue and then destained with acetic acid/methanol solution. Immunoblotting of proteins from polyacrylamide gel to a nitrocellulose filter was done according to the protocol of the ProBlot Western Blot AP System. For detection of the α-human atrial natriuretic peptide, the blots were incubated with 1:5000 dilution of primary antibody rabbit anti-hANP antiserum (Peninsula) and 1:1000 dilution of secondary antibody alkaline phosphatase (Promega).

RESULTS AND DISCUSSION

Construction of the fusion systems Figure 1 shows a schematic demonstrating the strategy of the construction of fusion gene. A 1.7 kb *Pst*I-*Stu*I fragment containing the L-asparaginase coding sequence was cleaved from plasmid pASN326 and isolated by agarose gel electrophoresis. The *Pst*I/*Stu*I double digested DNA fragment was blunt-ended by a trim-off reaction with T4 DNA polymerase and DNA polymerase I large (Klenow) fragment to trim off the 3' protruding end produced by the *Pst*I. The blunt ends of the fragment were inserted into M13mp10 at the *Sma*I site. The resulting recombinants were transformed into JM101 cells and the consequent recombinant transformants were selected by blue/white plaque color screening on λ plates containing IPTG and X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside). The white plaques were grown in 2X TY medium. The recombinant single-stranded DNA was isolated and sequenced by using M13 (-40) universal primer and five 9-mer primers.

Since there was not a restriction endonuclease site at the end of the signal sequence or at the stop codon, a double-site *in vitro* mutagenesis was performed to permit flexibility for this and other research. This mutation changed A to C and T to G (position 66 and 1051 of the L-asparaginase gene, respectively) to create an *Eag*I at the end of the signal sequence and a *Bst*XI at the stop codon. These changes did not alter the coding of the amino acid sequences. Mutagenic primers MASS and MTOE (Table 1) were annealed to the YN117 single-stranded DNA and the second strand DNA was synthesized with T7 DNA polymerase and 5-methyl-dCTP. After transfection with *E. coli* SDM, 24 different plaques were selected arbitrarily and analyzed by DNA sequencing. Six plaques had this double mutation (Figure 2).

Two oligonucleotides were chemically synthesized and phosphorylated at the 5' ends (5'-ATATCGAAGGTCGA-3' and 5'-AGCTTCGACCTTCGATATAAGT-3'). A linker sequence encoding the cleavage-recognition site (Ile-Glu-Gly-Arg) for protease factor Xa was formed by annealing these two phosphorylated synthetic oligonucleotides. This linker sequence contained a *Bst*XI end on its 5' terminus and a *Hind*III end on its 3' terminus that enabled the construction of fusion gene. The L-asparaginase gene on YN117 was digested with *Bst*XI and *Pst*I and the α-hANP gene on M13mp18 was digested with *Hind*III and *Pst*I to create compatible ends with the linker sequence. These three fragments--asparaginase gene, linker gene, and α-hANP gene--were fused by incubation at room temperature overnight and followed at 4°C for 6 h. The *Bst*XI is a restriction endonuclease which has an interrupted palindromic sequence (CCANNNNN/NTGG), so that the designed linker sequence would not

form a polymer more than a two mer in the ligation reaction. The ligation mixture was transformed to JM101 cells. The fusion gene then was constructed as asparaginase gene - linker gene - α-hANP gene on plasmid AN68.

Two constructions, one with and one without the signal sequence, were prepared to increase the fusion protein production and study the translocation of the native signal sequence (Figure 3). Plasmid AN68 was the DNA template for the polymerase chain reaction (PCR). For the construction with the signal sequence, two primers for PCR were used: primer PEWS, complementary to the 5' end of the signal sequence and also containing an extension on the 5' end to add an *Eco*RI site; and primer PPST, complementary to the 3' end of the fusion gene. The amplified DNA then was digested with *Eco*RI and *Pst*I restriction enzymes with the aid of six and eight extra base pairs hanging on the ends to generate compatible ends for cloning into an expression vector. The digested small ends were removed by applying to GENECLEAN II, and the purified large fusion gene was joined to vector pKK223-3 to construct an expression vector pCP32. The same procedures were applied to the construction of the fusion gene without the signal sequence except the primer in the PCR was PENL, instead of PEWS. Primer PENL, complementary to the sequence after the signal sequence, had an extension sequence for *Eco*RI site and ATG start codon. Together with primers PENL and PPST, a fusion gene without the signal sequence was created and inserted into vector pKK223-3 to construct an expression vector pCA919. A region of the L-asparaginase gene that included the active site was sequenced by Lofstrand Labs (Gaithersburg, MD) and found to be correct. (The gene corresponding to amino acids 115-147 was sequenced. The active site corresponds to amino acids 118-127).

Expression of the fusion proteins in *E. coli* The expression vectors pCA919 and pCP32 were transformed into the expression *E. coli* host strain JM105 cells to test the expression of the fusion proteins in the cytoplasm and periplasm, respectively, upon the IPTG induction. Cultures harboring the vectors pCA919 or pCP32 together with a culture harboring the expression vector pKK223-3 as control were grown at 37°C overnight. The overnight cultures were inoculated into fresh LB media and grown at 37°C to an optical density of 0.4 at 600 nm. Induction was turned on by the addition of 1 mM IPTG and incubation at a lower temperature of 30°C was continued for 3 h. The fusion proteins expressed by pCA919 or pCP32 under the regulation of *tac* promoter were produced in the cells. The cell crude extracts of each sample were resolved by SDS-PAGE and visualized by Coomassie brilliant blue staining (Figure 4). The molecular weights deduced from the amino acid sequence of fusion proteins with and without the signal sequence are approximately 41 kDa and 39 kDa, respectively. Comparison among the cell crude extracts revealed only a single band difference relative to the controls of host cell with pKK223-3. These specific bands had the size of 39 kDa which was predicated for the fusion proteins. These fusion protein bands were detected by immuno-enzymatic staining (Western blot) using rabbit anti-hANP serum and goat anti-rabbit IgG antibody-alkaline phosphatase conjugate staining. α-hANP was a portion of the fusion protein and could represent the location of the fusion protein with Western blotting. Figure 5 shows the result of Western blotting. Both bands from pCA919 or pCP32 at the size of 39 kDa reacted with anti-hANP serum. By using SDS-PAGE and immunoblotting, it was confirmed that these specific bands represented the position of the fusion proteins. Because of lacking the signal sequence, the fusion protein synthesized from vector pCA919 remained in the cytoplasm. A noteworthy difference between the fusion proteins from pCP32 and pCA919 is the fusion synthesized from pCP32 should have no

methionine at the N-terminal because it should have been cleaved during the translocation process but the fusion synthesized from pCA919 should retain the methionine as the start residue. This extra methionine on the asparaginase should have no significant influence on this research because the asparaginase would be separated after digestion with factor Xa. The expression levels of fusion proteins from both constructions were determined by scanning the SDS-polyacrylamide gel with a densitometer (Zeineh SLR). The expression of the fusion protein from the construction of pCA919 (no signal peptide) was 8.3% of total cell protein and the fusion protein from the construction of pCP32 (signal peptide present) was 3.6% of total cell protein.

The growth temperature effect One of the problems in producing eucaryotic proteins in *E. coli* is that the proteins may form insoluble inclusion bodies. The host cell, growth conditions, and conformation of the protein can affect the formation of inclusion bodies. Research on human interferon-α2 (IFN-α2) showed the effect of temperature on solubility appeared to be independent of the level of expression, and a comparison of the formation of inclusion bodies at the induction temperature 37°C or 30°C showed inclusion bodies reduced dramatically at 30°C (Schein *et al.*, 1988). By using the cell fractionation experiment, it was shown that the formation of inclusion bodies of the fusion protein from pCA919 was reduced dramatically at the induction temperature at 30°C compared with the inclusion bodies found upon incubation at 37°C (Figure 6).

Purification of the fusion protein A purification process was developed for the fusion protein expressed in the cytoplasm because of the higher level of expression compared to that in the periplasm. The cell lysate supernatant from sonication was subjected to affinity purification using L-asparagine immobilized on agarose. The reason to use L-asparagine as the ligand in the affinity purification was that the hydrolysis rate of L-asparagine was 20 times faster than that of D-asparagine for L-asparaginase from *Erwinia carotovora* (Howard *et al.*, 1972).

The flow rate used for loading of the cell lysate from 250 ml of broth was set at 0.1 ml/min. Unbound protein was extensively washed from the column with sodium phosphate buffer (10 mM, pH 7.5) at a flow rate of 0.6 ml/min. Elution was performed with the elution buffer containing 0.2 M sodium chloride at flow rate of 0.4 ml/min. The final eluents were collected in fractions. The individual fractions containing the L-asparaginase fusion protein were analyzed by SDS-PAGE. Coomassie blue staining of these fractions gave one dense band at the size of 39 kDa and a very faint band between 35 and 39 kDa; this indicates that the fusion protein was purified to very near homogeneity (Figure 7). The identity of the eluted protein was determined by Western blotting with controls of pKK223-3 cell crude extract and pCA919 cell crude extract. The purified protein could be recognized by the α-hANP antiserum.

The total protein concentration of affinity-purified fusion protein was measured using BCA protein assay reagents. The approximate binding capacity of the L-asparagine agarose was 300 μg of fusion protein per ml of gel. This value is probably lower than it should be because the gel has been reused many times.

Cation exchange chromatography was also selected to purify the fusion protein because both L-asparaginase and α-hANP have high isoelectric points (pI), which are 8.6 and 12.3 (Saito *et al.*, 1987), respectively. The pI of L-asparaginase fusion protein, estimated by a program "ISOELECTRIC" in University of Wisconsin Genetic Computer Program, is 9.47. This high pI enables cation exchange chromatography to be feasible.

The same operation was also applied to cation exchange chromatography. The 25 ml cell lysate in HEPES buffer was loaded to a CM-Sepharose Fast Flow column and washed with the same buffer to remove unbound proteins. The bound proteins were eluted with buffer with 0.2 M sodium chloride. The fusion protein was purified by this cation exchange chromatography to near homogeneity. Figure 8 shows the fractions analyzed by SDS-PAGE.

L-asparaginase has been proven to be proteolytically sensitive in a large-scale process (Lee *et al.*, 1989). Recently, a different purification procedure was used to isolate L-asparaginase from *Erwinia chrysanthemi* without evidence of proteolytic degradation of L-asparaginase (Goward *et al.*, 1992). The L-asparaginase was extracted at pH of 11.5 and extraneous protein was precipitated at pH 4.8. An S-Sepharose Fast Flow column was exploited to chromatograph the enzyme. The operation conditions in this process have circumvented the pH range 7.5 to 9.5, at which many proteases are active. This procedure is a possible alternative way to purify the L-asparaginase fusion protein on a large-scale.

CONCLUSION

An L-asparaginase fusion protein system was constructed to simplify the purification of eucaryotic peptides produced in *E. coli*. α-hANP was hybridized to the C-terminus of *Erwinia* L-asparaginase through a cleavable linker recognized by factor Xa. Two constructions were created: the fusion gene with the signal sequence, which could correctly target to the periplasm, and the fusion gene without the signal sequence, which would allow the fusion protein to remain in the cytoplasm. It was observed that the inclusion bodies were dramatically reduced when the induction temperature shifted from 37°C to 30°C. The expression of the fusion protein synthesized from the fusion gene with the signal sequence was 3.6% of total cell protein and from the fusion gene without the signal sequence was 8.3% of total cell protein. The fusion protein was purified to homogeneity or very near homogeneity under non-denaturing conditions by either affinity chromatography on immobilized L-asparagine or cation exchange chromatography with CM-Sepharose Fast Flow column.

The L-asparaginase fusion system has been found to have several advantages compared to other fusion systems that have been developed. The L-asparaginase fusion system uses L-asparagine as a ligand in affinity purification that is cheaper than the glutathione ligand used in the glutathione *S*-transferase (GST) fusion system. Its relatively high pI (9.47) allowed the fusion protein to be purified to homogeneity by cation exchange chromatography, a method that is cheaper than all types of affinity chromatography. Most of the fusion protein was soluble and thus could be purified under non-denaturing conditions. Even if a peptide joins to L-asparaginase and the fusion protein aggregates as inclusion bodies, it can not be purified by affinity chromatography but it is possible to purify the fusion protein by cation exchange chromatography after denaturation of the inclusion bodies. This feature makes this system attractive compared to the GST fusion system when the fusion protein is expressed in inclusion bodies, since the GST system does not work when inclusion bodies are formed. On the other hand, if the fusion protein can not be completely purified by a single-step cation exchange chromatography, the purification can still be simplified by passing cell lysate through cation exchange chromatography to remove most of proteins and facilitating the following affinity chromatography. This property is especially important in a large-scale purification.

REFERENCES

Chissoe, S.L., Wang, Y.F., Clifton, S.W., Ma, N.T., Sun, H.J., Lobsinger, J.S., Kenton, S.M., White, J.D. and Roe, B.A., Strategies for rapid and accurate DNA sequencing, Methods: A Companion to Methods in Enzymology, 3(1), 55-65 (1991).

de Bold, A.J., Atrial natriuretic factor: a hormone produced by the heart, Science, 230, 767-770 (1985).

Gilbert, H.J., Blazek, R., Bullman, H.M.S. and Minton, N.P., Cloning and expression of *Erwinia chrysanthemi* asparaginase gene in *Escherichia coli* and *Erwinia carotovora*, J. of General Microbiology, 132, 151-160 (1986).

Goward, C.R., Stevens, G.B., Tattersall, R. and Atkinson, T., Rapid large-scale preparation of recombinant *Erwinia chrysanthemi* L-asparaginase, Bioseparation, 2, 335-341 (1992).

Howard, J.B., and Carpenter, F.H., L-asparaginase from *Erwinia carotovora*, J. Biol. Chem., 247, 1020-1030 (1972).

Laemmli, U.K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature, 227, 680-685 (1970).

Lee, S.M., Wroble, M.H. and Ross, J.T., L-asparaginase from *Erwinia carotovora*: an improved recovery and purification process using affinity chromatography, Applied Biochem. and Biotech., 22, 1-11 (1989).

Löwenadler, B., Jansson B., Paleus, S., Holmgren, E., Nilsson, B., Moks, T., Palm, G., Josephson, S., Philipson, L. and Uhlén M., A gene fusion system for generating antibodies against short peptides, Gene, 58, 87-97 (1987).

Maina, C.V., Riggs, P.D., Grandea, A.G., Slatko, B.E., Moran, L.S., Tagliamonte, J.A., McReynolds, L.A. and Guan, C., An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein, Gene, 74, 365-373 (1988).

Marston, F.A.O., The purification of eucaryotic polypeptides synthesized in *Escherichia coli*, Biochem. J., 240, 1-12 (1986).

Minton, N.P., Bullman, H.M.S., Scawen, M.D., Atkinson, T. and Gilbert, H.J., Nucleotide sequence of the *Erwinia chrysanthemi* NCPPB1066 L-asparaginase gene, Gene, 46, 25-35 (1986).

Neu, H.C. and Heppel, L.A., The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts, J. of Biol. Chem., 240, 3685-3692 (1965).

Offensperger, W., Wahl, S., Neurath, A.R., Price, P., Strick, N., Kent, S.B.H., Christman, J.K. and Acs, G., Expression in *Escherichia coli* of a cloned DNA sequence encoding the pre-S2 region of hepatitis B virus, Proc. Natl. Acac. Sci. USA, 82, 7540-7544 (1985).

Oka, T., Sakamoto, S., Miyoshi, K.I., Fuwa, T., Yoda, K., Yamasaki, M., Tamura, G. and Miyake, T., Synthesis and secretion of human epidermal growth factor by *Escherichia coli*, Proc. Natl. Acad., Sci. USA, 82, 7212-7216 (1985).

Saito, Y., Ishii, Y., Koyama, S., Tsuji, K., Yamada, H., Terai, T., Kobayashi, M., Ono, T., Niwa, M.,Ueda, I. and Kikuchi, H., Bacterial synthesis of recombinant α-human atrial natriuretic polypeptide, J. Biochem., 102, 111-122 (1987).

Schein, C.H. and Noteborn, M.H.M., Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature, Bio/Technology, 6, 291-294 (1988).

Smith, D.B. and Johnson, K.S., Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase, Gene, 67, 31-40 (1988).
Uhlén M. and Moks, T., Gene fusions for purpose of expression: an introduction, Methods in Enzymol., 185, 129-143 (1990).
Vandeyar, M.A., Weiner, M.P., Hutton, C.J. and Batt, C.A., A simple and rapid method for the selection of oligodeoxynucleotide-directed mutants, Gene, 65, 129-133 (1988).
Wriston J.C., Asparaginase, Methods in Enzymology, 18, 732-742 (1970).

Table 1 Oligonucleotide sequences used in DNA sequencing, *in vitro* mutagenesis, and polymerase chain reaction

DNA sequencing primers:

M13 (-40) universal primer: 5'-GTTTTCCCAGTCACGAC-3'

P2: 5'-GTACGGCTC-3'

P3: 5'-TGCGAGTCA-3'

P4: 5'-AATGTCGAC-3'

P5: 5'-AGCACTGAT-3'

P6: 5'-AGCGTATCC-3'

Mutagenic primers:

MASS: 5'-AGTTTATCGGCCGCACTC-3'

MTOE: 5'-AGACAATCCATAAGTATG-3'

Polymerase chain reaction primers:

PEWS:

5'-TCACCGGAATTCATGGAAAGATGGTTTAAATCTCTG-3'

PENL:

5'-TCACCGGAATTCATGGCCGATAAACTGCCCAATATC-3'

PPST: 5'-AGCTTGGGCTGCAGCTATTAG-3'

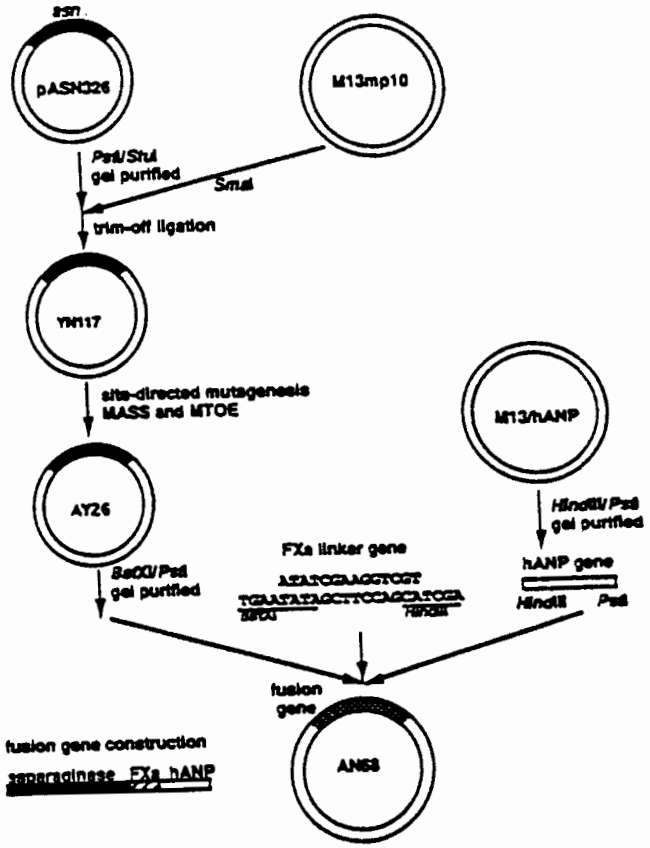


Figure 1 Strategy for the construction of the fusion gene

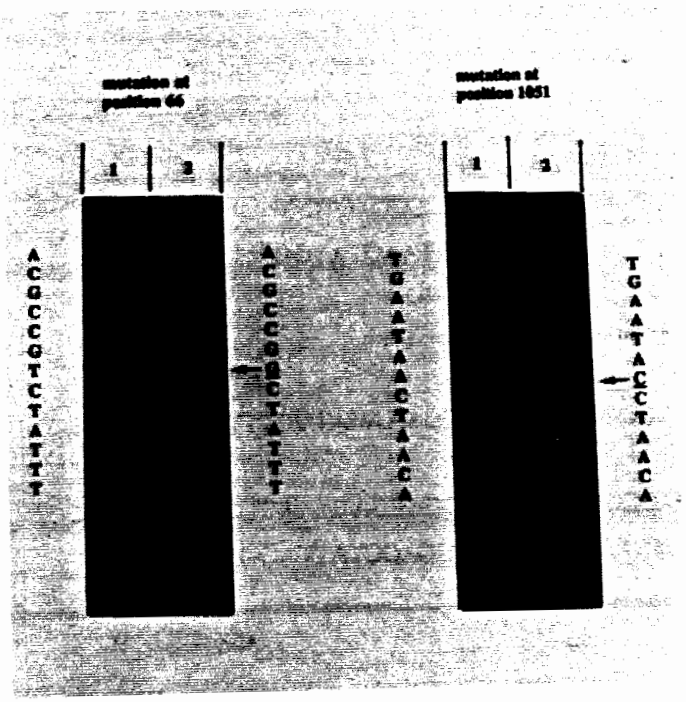


Figure 2 DNA sequencing of the mutated region of the L-asparaginase gene
The order of bases for each clone is ACGT. The sequences shown are those of the complementary strands in the vicinity of positions of 66 (A→C) and 1051 (T→G) (arrowed). Lane 1, DNA sequences of wild type clones in the region of the mutation; lane 2, DNA sequences of mutant clones in the region of the mutation.

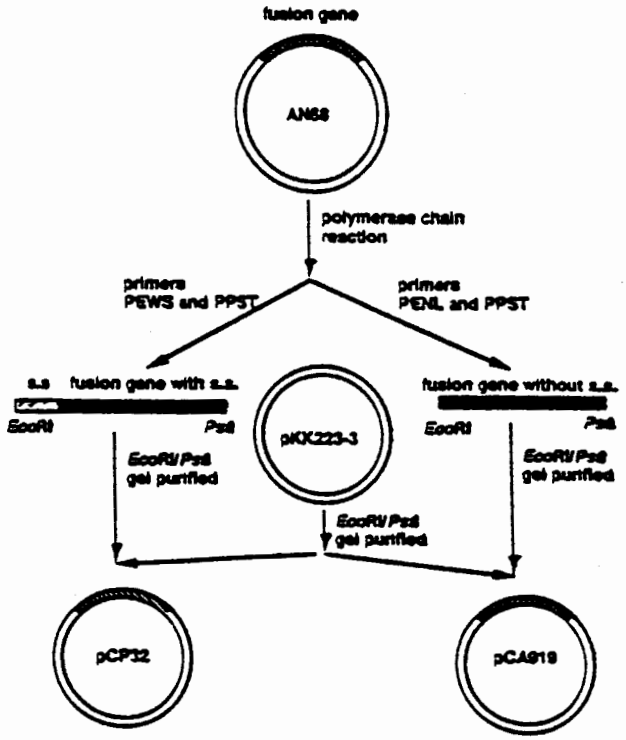


Figure 3 Strategy for the construction of the fusion genes with and without the signal sequence on the expression vectors

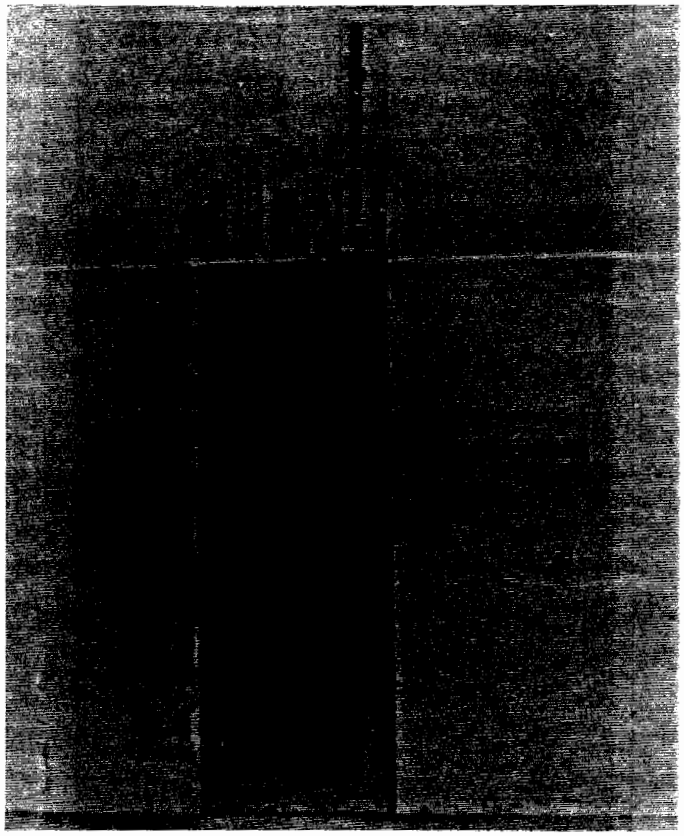


Figure 4 Analysis of the fusion proteins expression by SDS-PAGE and Coomassie blue staining

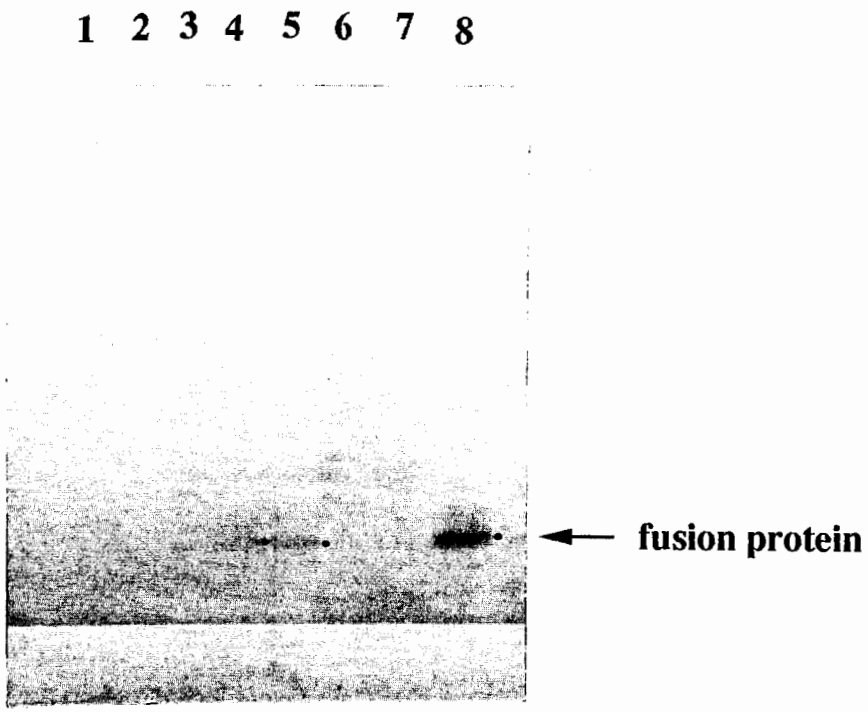


Figure 5 Western blot analysis of the fusion proteins expression
Lane 1, prestained marker, lanes 2 and 6, JM105 crude extracts; lanes 3 and 7, cell crude extracts containing pKK223-3; lanes 4 and 5, cell crude extracts containing pCP32 and lane 8, cell crude extracts containing pCA919.

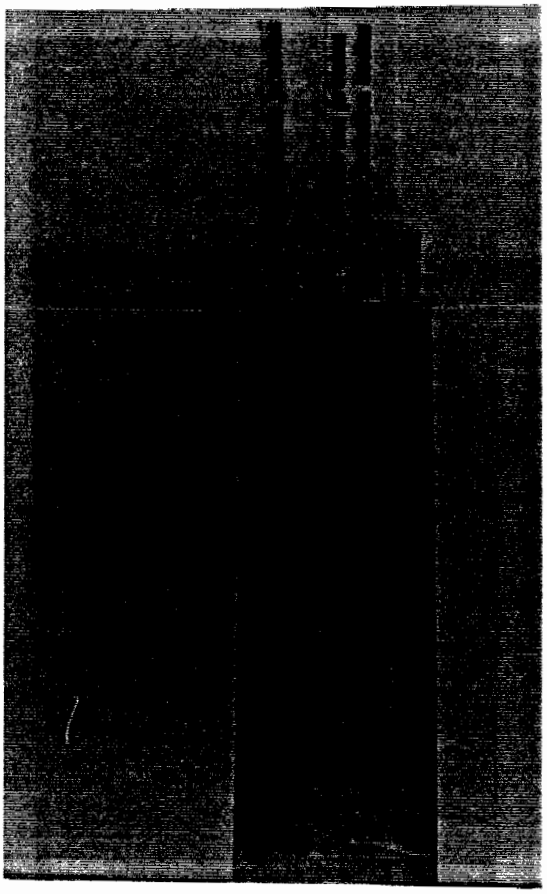


Figure 6 SDS-PAGE analysis of the formation of insoluble inclusion bodies at the induction temperature 37°C and 30°C

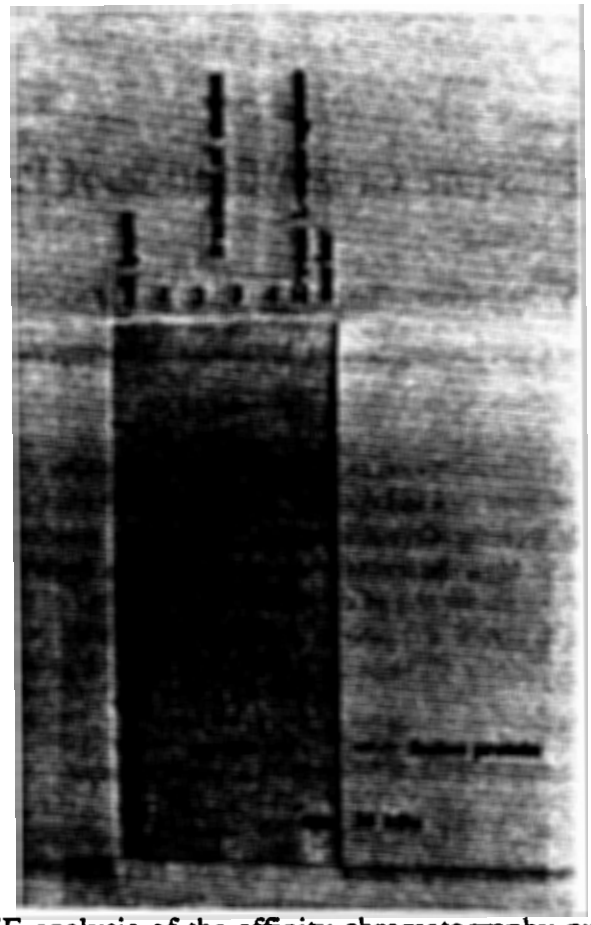


Figure 7 SDS-PAGE analysis of the affinity chromatography purified fusion protein with Coomassie blue staining
Lanes 1 to 4 are the fractions of elution. The arrow designates the position of the fusion proteins.

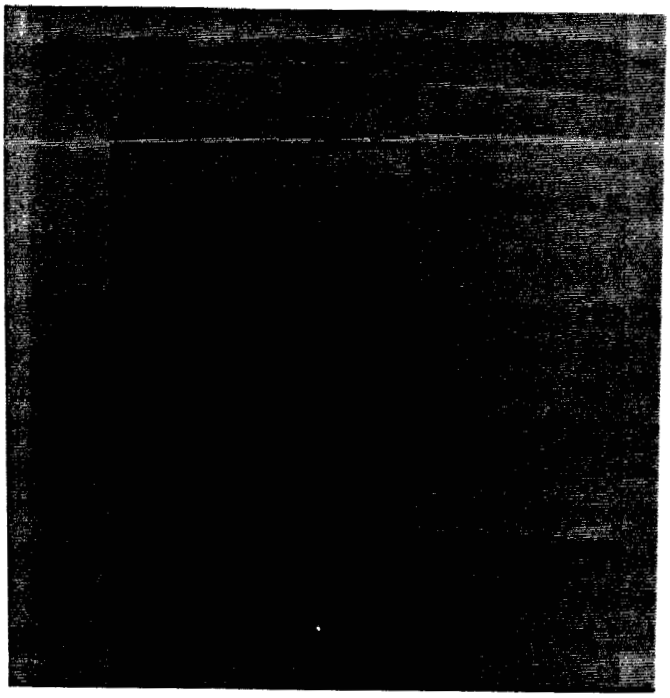


Figure 8 SDS-PAGE analysis of the cation exchange chromatography purified the fusion protein with Coomassie blue staining
Lanes 1 to 6 are the fraction of elution. The fusion protein is arrowed.

High Pressure Crystallization of Proteins

Mungara V. Saikumar, Charles E. Glatz and Maurice A. Larson

Department of Chemical Engineering, Iowa State University, Ames, IA 50011

Several methods for crystallization of biological macromolecules have been proposed in the past to obtain crystals of good size and quality. The effect of pressure on protein crystallization has been explored only recently. Crystallization of baro-stable proteins was undertaken under pressures as high as 150 Mpa. Preliminary experiments using lysozyme as a model protein showed that pressure increases the rate of crystallization. The effect of the time of pressurization on the yield of protein in crystal and activity recovered at different pressures showed only 0.1 MPa pressure to be significantly different from the others for both protein yield and activity recovered in crystals.

Introduction

Many complex molecules produced by the biotechnology and pharmaceutical industries are supplied in solid, often crystalline form. The final solid product characteristics of interest typically include morphology, particle size distribution, purity and stability - characteristics governed by crystallization. Proteins, which are high molecular weight macromolecules, are difficult to crystallize. A large number of techniques have been developed for protein crystallization and growth in the past [1]. However, most of these concentrate on obtaining protein crystals of sufficient size for X-ray crystallography. In our work high pressures were employed to crystallize lysozyme, emphasizing increasing the rate of crystallization rather than the crystal size.

High pressure protein crystallization

Pressure effects on protein conformation and enzyme kinetics have been studied in the past [2,3,4]. Reversible and irreversible changes in conformation of the proteins when subjected to high pressures for extended periods of time were observed. Lysozyme and amylase were stable up to pressures as high as 100 MPa. Pepsin and invertase were stable up to pressures as high as 5000 MPa [5]. Lysozyme, a model protein for most protein crystallization studies was used in our work.

Pressure was first used to crystallize glucose isomerase by Visuri et al. [6]. It was seen that glucose isomerase crystallized orders of magnitude faster at 200 MPa than it did at atmospheric pressure (0.1 MPa). The enzyme crystallized in 15 minutes at high pressures as opposed to hours to days at atmospheric pressure. The rate of crystallization increased with increasing pressures and at 200 MPa was approximately 17 times higher than at 0.1 MPa. It took about 6 hours to reach the equilibrium solubility in stirred batches at 0.1 MPa and more than 24 hours in unstirred batches for the onset of crystallization. At a pressure of 200 MPa nucleation

was seen to begin in less than 2 minutes. At higher pressures (300 MPa) there was a formation of amorphous and insoluble precipitate rather than crystals and hence Visuri et al. did not go beyond 200 MPa. The authors however, presented no hypothesis for the dramatic increase in the rate of crystallization.

The increase in the rate of crystallization can be explained by treating the crystallization process as a reaction, the rate constant dependency on pressure is as follows

$$\left(\frac{\partial \ln(k)}{\partial p}\right)_T = -\frac{\Delta V^*}{RT}$$

where k=rate constant, R=gas constant, T=temperature, ΔV^*=change in volume between transition state and reactant, P=pressure. Thus a negative activation volume for nucleation or growth coupled with increase in pressure would increase the rate constant. Since the compressibilities of the enzymes are very low, large pressures are required to obtain a significant increase in the rate constant and hence in the rate of crystallization.

Materials and methods

Hen egg white lysozyme purchased from Sigma Chemical Co. (Grade 1, 95% protein) was used directly without further purification. Lysozyme was dissolved in 50mM sodium acetate, pH 4.5, buffer to make up a 50 mg/ml protein stock solution. Sodium chloride stock solution, 5.2M, was made by dissolving sodium chloride in the same 50mM acetate buffer. Both the stock solutions were filtered using a 0.45μm Whatman filter paper. The stock solutions were added to the crystallization cell to make a solution of 40 mg/ml of lysozyme and 0.8M sodium chloride with a total volume of 1 ml. The cell used was the tube end of a disposable polyethylene transfer pipette which was heat sealed at one end. After it was filled with the crystallization solution the other end was heat sealed. It was then placed in a pressure bomb and the bomb placed in the pressure apparatus shown Fig. 1. System pressure was then raised to the desired level. After pressurizing for a fixed time, the cell was removed from the high pressure bomb and the contents were allowed to crystallize for a total of 36 hours from the time the sample was prepared, with all the time after the release of pressure being at atmospheric pressure. Crystals were separated by centrifugation and were redissolved in the acetate buffer to determine the protein and activity content of the crystals. Total protein concentration was obtained using OD at 280, and the enzymatic activity was obtained by measuring the change in OD at 540nm of a 50 mg% solution of *Micrococcus lysodeikticus* per unit time.

Results and discussion

An unreplicated factorial experiment was performed at two different pressures (100 and 150 MPa), pressurized for two different times (2 and 8 hours) with two different initial con-

centrations of the protein (50 and 80 mg/ml). The 80 mg/ml enzyme solution was obtained by concentrating the solution by ultrafiltration (1000 MWCO). Table 1 summarizes the results of this experiment. It can be seen that the yields of proteins as crystals was higher for higher initial concentration, longer time of pressurization and higher pressures with some inconsistencies. In all cases protein yields were higher when pressure was applied.

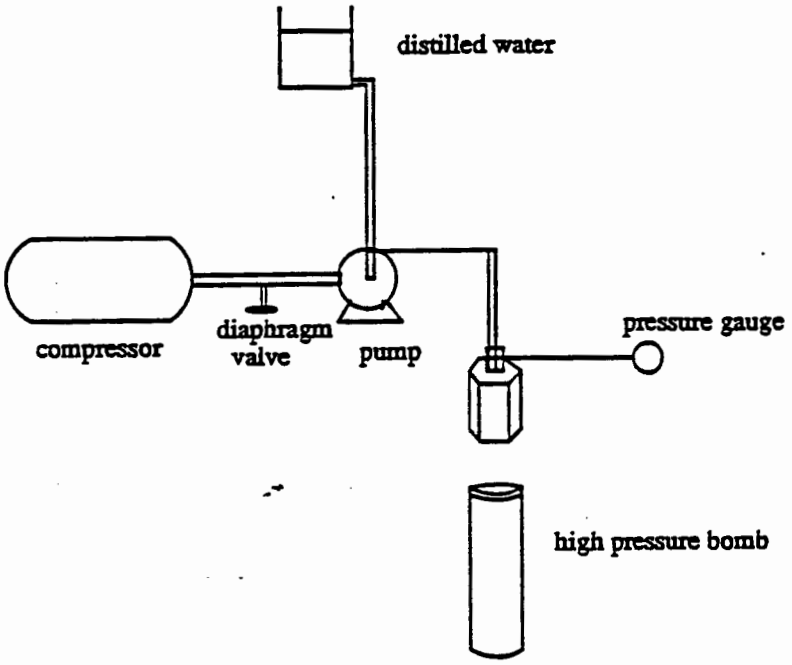


Fig. 1. Schematic of high pressure equipment

Table 1 Effect of pressure and time of exposure to pressure on the amount of total protein in crystalline phase. Total time of crystallization is 36 hours.

Time (h)	Concentration (mg/ml)	Pressure (MPa)	Protein Yield (%)
8.0	80.0	150	84
8.0	80.0	100	64
8.0	50.0	150	89
8.0	50.0	100	56
2.0	80.0	150	67
2.0	80.0	100	79
2.0	50.0	150	52
2.0	50.0	100	42
	80.0	0.1	49
	50.0	0.1	33

Figure 2 shows the change in concentration of the crystallizing solution with time (initial concentration of 37.2 mg/ml) when exposed to 100 MPa for 2 hours and then left at atmospheric pressure. Four samples were subjected to pressure and were analyzed after 2, 12, 24 and 56 hours. Crystals were isolated in the first sample as soon as the pressure was released. The others were left at atmospheric pressure for 10, 22, 54 hours respectively. It can be seen that the sample exposed to pressure crystallized faster than the one not subjected to pressure.

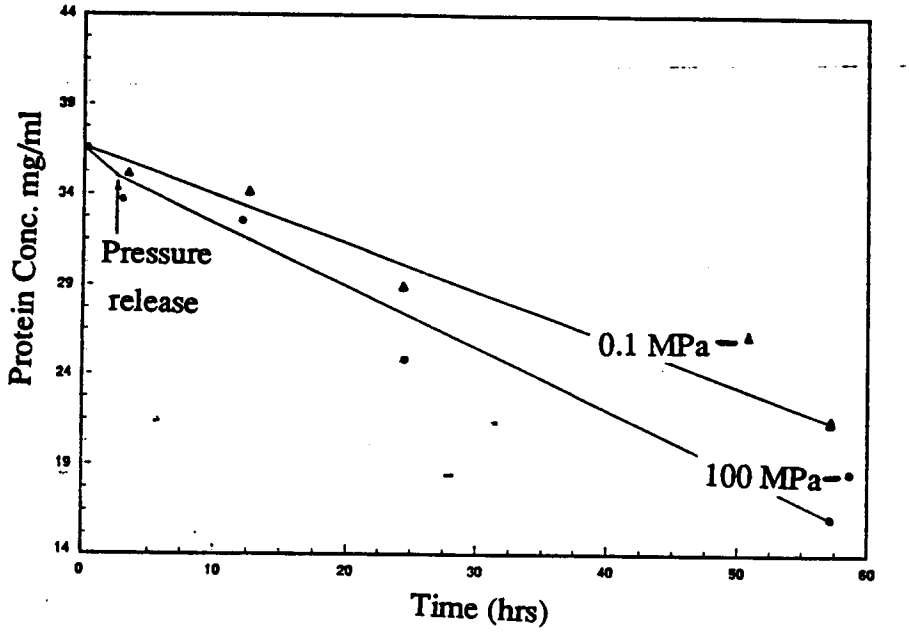


Fig 2. Decrease in protein concentration in solution with time

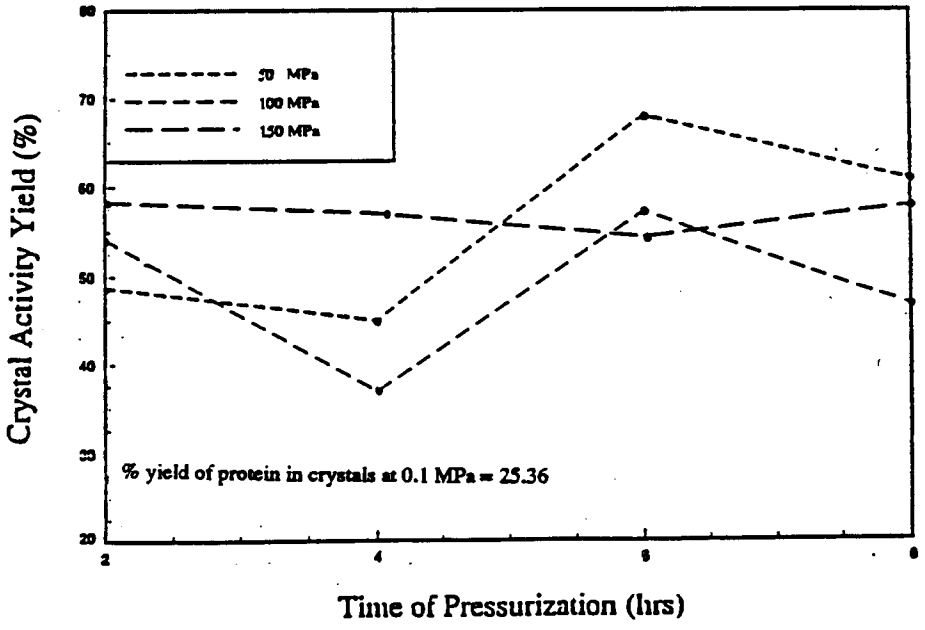


Fig. 3. Effect of time of pressurization on the percentage yield of protein in crystals

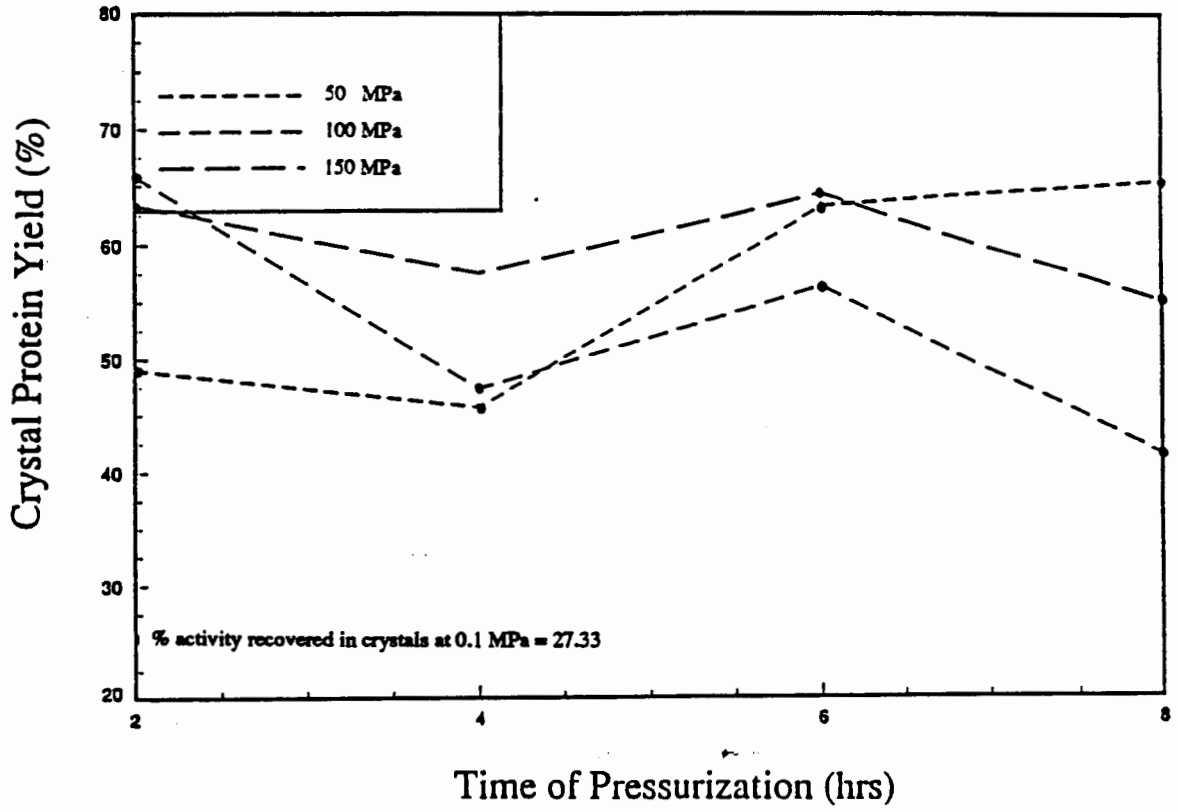


Fig. 4. Effect of time of pressurization on the Percentage of activity recovered in crystals

Figures 3 and 4 were obtained from a completely randomized 2x4 factorial design with pressure and time as the two factors. Pressure levels were with 4 levels in each of them, 0.1, 50, 100 and 150 MPa, times of 2, 4, 6 and 8 hours at pressure. Figure 3 shows the effect of the time of pressurization on the yield of crystalline protein at the four different pressures. It is seen that the protein yield of all the samples subjected to pressure was significantly higher than the one at atmospheric pressure. Figure 4 shows the percent activity yield in crystals of the same samples (average value of the two replicates was plotted in both cases). Statistical analysis of the data was done using SAS showed the effect of pressure on the crystal protein and activity yield is significant at the 95% level of confidence. The effect of time of pressurization on the crystal activity yield was significant at 95% level but on the crystal protein yield was significant only at 80% level of confidence. Paired comparisons of the time effects showed only 0.1 MPa pressure to be different. (The effects of the three pressures 50, 100 and 150 MPa were not significantly different from each other). The two plots do not show any definite trends with respect to the different levels of time of pressurization and levels of pressure. This could be due to the large variability in the data. Further experimentation is being directed to reduce this variability by slightly modifying the method.

Having seen that pressure had a significant effect on the rate of crystallization, experiments were then directed to see whether pressure was effecting nucleation and/or growth. To address this question a sample was left under pressure (50 MPa) for 36 hours and yield was compared with the earlier samples as shown in Table 2.

Table 2 Effect of Pressure on nucleation and/or growth

Sample No.	Pressure (MPa)	Time of Pressurization (h)	Time from preparation before Analysis	Protein yield in crystals (%)
1	50 MPa	36	36 hrs	35.0
2	0.1 MPa		8 hrs	15.0
3	50 MPa	8	8 hrs	17.0
4	50 MPa	8	36 hrs	65.0
5	0.1 MPa		36 hrs	25.0

The column 4 in Table 2 gives the total time of crystallization before centrifuging to separate out the crystals. The protein yield of a sample pressurized for 36 hours (sample 1) is lower than the one under pressure for 8 hours and left at atmospheric pressure for the remaining of the 36 hours (sample 4). Comparing the yields of these two and with sample 5 it can be seen that pressure is aiding nucleation more than growth. The larger number of nuclei formed under pressure grow to give higher yields in the case of sample 4. The yields obtained in 2 and 3 support this hypothesis. It can thus be concluded that high pressures are aiding nucleation more than growth.

Conclusion

It can be seen that pressure is increasing the rate of crystallization of lysozyme, however, not as dramatically as had been reported for glucose isomerase. The crystal protein yield was significantly higher at high pressures. However, the effects due to the different levels of pressure and the different times of pressurization were not distinguishable. The results obtained did not show any definite trends with respect to the different levels of pressure or time. Additional work is being directed to reduce the variability and see whether this lack of a definitive trend is due to the variability in the data or whether there is another factor which is being neglected. Pressure seems to be influencing nucleation more than growth. Additional work is being directed to verify this finding.

Acknowledgment

The authors thank Dr. Bert E. Nordlie, Professor, Geological and Atmospheric Sciences Department, Iowa State University for permitting the use of high pressure equipment in his lab for the protein crystallization studies.

References

1. McPherson, A. 1982. *Preparation and Analysis of Protein Crystals*. John Wiley & Sons, N.Y.

2. Gekko, K., and Noguchi, H. 1979. *J. Phys. Chem.* 83: 2706-2714.

3. Chryssomallis, G.S., Torgerson, P.M., Drickamer, H.G. and Weber, G. 1981. *Biochemistry* **20**: 3955-3959.

4. Jason, E.M. Jr., Dow, R.B. and Anderson, A.K. 1940 *J. Biol. Chem.* **135**: 697-705.

5. Jaenicke, R. 1981. *Ann. Rev. Biophys. Bioeng.* **10**: 1-67.

6. Visuri, K., Kaipainen, E., Kivimaki, J., Niemi, H., Leisola, M. and Polosaari, S. 1990. *Bio/Technol.* 8: 547-549.

Structure/Function Relationships in the Catalytic and Starch Binding Domains of Glucoamylase

Pedro M. Coutinho[1], Clark Ford[2], Peter J. Reilly[1]
Departments of Chemical Engineering[1] and Food Science and Human Nutrition[2]
Iowa State University, Ames, IA 50011

INTRODUCTION

Glucoamylase (E.C. 3.2.1.3, Glucan 1,4-α-glucosidase) is an industrial enzyme that hydrolyzes terminal 1,4-linked α-D-glucose residues successively from the non-reducing ends of poly- and oligo-sacharide chains with the release of ß-D-glucose (Pazur and Ando, 1959). Most forms of the enzyme can also hydrolyze 1,6-α-D-glucosidic bonds at a lower rate. Purified forms of glucoamylase are used in the degradation of dextrins to produce glucose, an intermediate step in the production of high-fructose corn syrups and ethanol. Microorganisms carrying the glucoamylase gene are used in food and alcoholic fermentation (Sakaguchi *et al.*, 1992).

Fifteen complete protein sequences of filamentous fungi, yeast and bacterial origin have been reported (Boël *et al.*, 1984; Nunberg *et al.*, 1984; Yamashita *et al.*, 1985; Ashikari *et al.*, 1986; Itoh *et al.*, 1987; Yamashita *et al.*, 1987; Hayashida *et al.*, 1989; Dohmen *et al.*, 1990; Shibuya *et al.*,1990; Hata *et al.*, 1991; Hostinová *et al.*, 1991; Aleshin *et al.*, 1992; Joutsjoki and Torkkeli, 1992; Ohnishi *et al.*, 1992; Berka *et al.*, 1993). Recently, the model of the crystal structure of the catalytic domain of glucoamylase from *Aspergillus awamori var. X100* has been obtained in its native form (Aleshin *et al.*, 1992) and in complexes with 1-deoxinojirimycin (Harris *et al.*, 1992) and acarbose (Aleshin, 1992). The catalytic domain is composed of thirteen α-helices, twelve forming an α/α barrel. The active site lies in a cavity in the center of the barrel.

Genetic engineering of glucoamylase for characterization (Itoh *et al.*, 1989; Sierks *et al.*, 1989; Sierks *et al.*, 1990; Sierks *et al.*, 1993; Sierks and Svensson, 1993) and modification of the enzyme to improve its industrial properties use are on course to increase selectivity towards glucose production (Sierks and Svensson, 1992a), the pH optimum (Bakir *et al.*, 1993) and the thermal stability (Chen *et al.*, 1993) of the enzyme. Information taken from tridimensional models of enzymes can be of extreme use, as they have been in modifying subtilisin.

Peptide mapping, gene sequencing, glycosylation evaluation, and deletion studies enabled the identification of three different regions in fungal glucoamylases: a catalytic domain (CD), an O-glycosylated region, and a starch-binding domain (SBD). Homology studies using fungal and yeast glucoamylase primary sequences identified several conserved segments, four by Tanaka *et al.* (1986) and five by Itoh *et al.* (1987), in the CD. These five conserved segments, designated S1 to S5, compose the loops surrounding the active site (Aleshin *et al.*, 1992).

Homology studies with other starch-degrading enzymes presented a conserved area in the C-terminal region common to fungal glucoamylases (Svensson *et al.*, 1989), and the presence of this region of glucoamylase in the binding to raw starch has been fully characterized (Dalmia and Nikolov, 1991). The homology of the SBD region to the C-terminal region of cyclodextrin glucosyltransferases (CGTases) and the recent obtention of the tridimensional models of CGTases from *Bacillus stearothermophilus* (Kubota *et al.*, 1990; Kubota *et al.*, 1991) and *Bacillus circulans no. 8* (Klein and Schulz, 1991), in the native state and in complexes with substrates allows the evaluation of some of the features of the SBD of fungal glucoamylases. This C-terminal domain of CGTases is composed of six to eight ß-sheets.

Glucoamylases have several features in common with cellulases. Cellulases can also be multidomain enzymes, cellulase binding domains having been identified in bacteria and in fungi (Bégin *et al.*, 1992; Wood, 1992) . These domains can be connected to the catalytic domain by regions rich in serine, tyrosine and proline residues corresponding in fungi to potentially O-

glycosylated regions. Recently, a bacterial cellulase presenting an α/α barrel structure has been reported (Juy *et al.*, 1992).

Families of cellulases and α-amylases have been defined by sequence alignment and hydrophobic cluster analysis. The cellulase families defined (Henrissat and Mornon, 1990) present different structures performing the catalytic activity, while all α-amylase families share the same structural features (Raimbaud *et al.*, 1989).

The modelling of substrates in the active site has been based on complexes with the substrate in α-amylase (Matsuura *et al.*, 1984) and CGTase (Klein *et al.*, 1992), on complexes with inhibitors for cellulase (Juy *et al.*, 1992) or simply by evaluation of the native crystal structure for α-amylase (Petkhov and Mazur, 1992). The evaluation of the interactions of residues in the active site and the subtrate are of great importance in the planning of changes affecting the catalytic activity.

EXPERIMENTAL METHODS

Multisequence alignment

The primary sequences of glucoamylases and of a CGTase used in this study have been retrieved or constructed using the **GCG** package, version 7.2-UNIX (Devereux *et al.*, 1984; Genetics Computer Group, 1991) on a DEC 2100 workstation. The glucoamylase sequences are: *Aspergillus awamori var. X100* (Aleshin *et al.*, 1992), *Aspergillus awamori var. kawachi* (Hayashida *et al.*, 1989), *Aspergillus shirousami* (Shibuya *et al.*, 1990), *Aspergillus niger* (Boël *et al.*, 1984), *Aspergillus awamori* (Nunberg *et al.*, 1984), *Aspergillus oryzae* (Hata *et al.*, 1991), *Humicola grisea var. thermoidea* (Berka *et al.*, 1993), *Hormoconis resinae* (Joutsjoki and Torkkeli, 1992), *Rhizopus oryzae* (Ashikari *et al.*, 1986), *Saccharamycopsis fibuligera* (Itoh *et al.*, 1987), *Saccharomycopsis fibuligera KZ* (Hostinová *et al.*, 1991), *Saccharomyces cerevisae* (Yamashita *et al.*, 1987), *Saccharomyces diastaticus* (Yamashita *et al.*, 1985), *Schwanniomyces occidentalis* (Dohmen *et al.*, 1990), *Clostridium sp.* (Ohnishi *et al.*, 1992). The CGTase sequence used is *Bacillus circulans no. 8* (Nitschke *et al.*, 1990). The primary sequences of the *A. niger* and *A. awamori* glucoamylases are the same.

Sequences were compared using the program **lineup** included in the **GCG** package that performs a multisequence progressive alignement (Feng and Doolitle, 1987) using as parameters a gap weight penalty of 1.75 and a gap length weight of 0.075. Preliminary results confirmed the existence of non-sequence homology between the *Schwanniomyces occidentalis* sequence and the remaining sequences (Dohmen *et al.*, 1990) and so the sequence was eliminated from further studies. The procedure was used in the alignment of fungal and yeast glucoamylase sequences and of the sequences corresponding to the SBD of fungal glucoamylases and the C-terminal domain of the *B. circulans* CGTase. The alignment of the glucoamylase from *Clostridium sp.* (Ohnishi *et al.*, 1992) was performed manually to improve the homology in the conserved segments due to the great differences with the remaining sequences that **lineup** could not take into account.

The distances between the fungal and yeast glucoamylase sequences were evaluated using program **distances** of the **GCG** package, based on the number of identical residues found between every two sequences. Families of glucoamylases were defined as having an identity of 50% or more of the residues after alignment of the fungal and yeast glucoamylases.

Modelling of maltose in the active site

The structure of the catalytic domain of glucamylase from *Aspergillus awamori var. X100* (Aleshin *et al.*, 1992), (**1gly** in the Brookhaven Protein Data Bank) was used in the modelling of α-methyl maltose in the active site. This study was achieved using program **QUANTA** (Molecular Simulations, Inc) on a Silicon Graphics 4D/35 workstation.

An α-methyl maltose molecule was constructed using **ChemNote** and changed in **QUANTA** to the conformation presented by more than 80 % of the molecules in aqueous solution

(Lemieux, 1992) — (ϕ / φ) = -25°/-15°. This conformer was manually superimposed to the 1-deoxynojirimycin molecule found in subsite 1 in the respective model of the crystal structure with glucoamylase (Harris *et al.*, 1993). The coordinates for α-methyl maltose (without hydrogen atoms) obtained were added to the **1gly** structure. A network of hydrogen bonds was evaluated for the glucoamylase-maltose complex based on visual analysis using **QUANTA.**

Evaluation of functionally important residues in glucoamylases

The residues involved in the binding of maltose in the active site indicated by the α-methyl maltose modelling were evaluated for all structures based on the multisequence alignment.

Residues involved in the binding of the SBD from fungal glucoamylases to maltooligosaccharides were evaluated based on residues interacting with maltose in the C-terminal domain of CGTase from *Bacillus stearothermophilus* in the crystal structure of the respective complex (Kubota *et al.*, 1990; Kubota *et al.*, 1991).

Convergent functions for some residues were found based on the tridimensional structure **1gly**, the C-alpha structure of the C-terminal domain of *Bacillus circulans no. 8* CGTase (Schulz, 1993) and the results of the multisequence alignment.

RESULTS AND DISCUSSION

Evaluation of glucoamylase families

Five families of structurally related glucoamylases were defined, two of filamentous fungi, two of yeast, and one of bacteria (Table 1).

The *Aspergillus* family comprises the *Aspergillus*, *Humicola* and *Hormoconis* glucoamylases. These seven sequences have a highly homologous primary structure and present as a common characteristic the same disulfide bridges in the CD as present in the crystal model. All, excluding the sequence **1gly** (a proteolytic product), present a C-terminal SBD. The linker region between the two domains, highly O-glycosylated in some of the sequences, presents a great variability.

The *Rhizopus* family contains only one sequence. This fungal sequence is at the most 30% identical to the members of the *Aspergillus* family and does not have the same disulfide bridges in the CD. A N-terminal SBD can be found linked to the CD by an O-glycosylated region.

The *Saccharomycopsis* family consists of two very similar sequences. The difference of only seven residues in the sequence of these two glucoamylases leads to surprisingly different behaviors under different pH conditions and temperatures (Gašperík and Hostinová, 1993). These glucoamylases do not have an SBD.

The two *Saccharomyces* sequences are closely related (Yamashita *et al.*, 1987). These sequences have a different N-terminal region and apparently no SBD. The N-terminal region of *S. diastaticus* is the result of a fusion of a glucoamylase gene similar to the one from *S. cerevisae* with additional gene sequence, part of which encodes an O-glycosylated region that functions in secretion (Yamashita, 1989). These regions also have some differences from the area of the conserved segment S5 to the C-terminus.

The only sequenced bacterial glucoamylase forms the *Clostridium* family. This sequence has a long N-terminal region with a Cys residue at the N-terminus that binds the enzyme to the cell wall (Onhishi *et al.*, 1991). This enzyme presents several insertions in the loops corresponding to the conserved segments but careful observation reveals its relationship to the remaining sequences.

The glucoamylase from *Schwannomyces occidentalis*, not included in this study, displays no clear homology to any of the other glucoamylase sequences. This seems to be a case of an enzyme belonging to the family of sugar-related enzymes that includes intestinal sucrase-

isomaltase and lysosomal α-glucosidase (Naim *et al.*, 1991), that acquired the catalytic activity that defines a glucoamylase through convergent evolution.

Modelling of α-methyl maltose in the active site

The analysis of the model of binding α-methyl maltose showed that this molecule takes the approximate position and conformation of the first two residues of acarbose in its complex with glucoamylase (Aleshin, 1992). The residues involved in the hydrogen bond with maltose are shown in Figure 1.

Studies with analogues of maltose had previously shown that the hydroxyl groups 4'-OH, 6'-OH, and 3-OH were essential for the substrate to be hydrolyzed (Bock and Pedersen, 1987). The hydroxyl groups 4'-OH and 6'-OH interact with Asp55, and 6'-OH forms a hydrogen bond with the catalytic water molecule. The hydroxyl group 3-OH makes a hydrogen bond with Arg305 and with the backbone oxygen of Trp178. The catalytic acid, Glu179 (Sierks *et al.*, 1990), lies close to the aglycon, with the catalytic base, Glu400 (Harris *et al.*, 1993) found in the opposite site, interacting with the catalytic water molecule. This spatial relation agrees with the proposed catalytic mechanisms (Harris *et al.*, 1993) and with the liberation of ß-glucose after hydrolysis of the α-glycosidic bond. The majority of the residues shown in Figure 1 are present in the conserved region segments.

Active site residues in all glucoamylases

The multisequence alignment in the area of the five conserved segments S1 to S5 was analyzed to verify the importance of the identified residues in glucoamylase catalytic action. The result of this study is displayed in Figure 2.

In S1 the four identified residues, Tyr48, Trp52, Arg54 and Asp55, in *A. niger* / *A. awamori* numbering, that are found at the end of the loop and are conserved in all the thirteen sequences, except that Arg54 is a Pro residue in *H. grisea var. thermoidea* in what could be a sequencing error. Modifications in *S. fibuligera* equivalent to W52R and W52L, to R54K and R54T and to D55N and D55Y lead to total loss of activity (Itoh *et al.*, 1989), and the modification D55G in *A. niger* lead to a significant decrease of of enzyme activity (Sierks and Svensson, 1993).

The S2 region presents four residues, Lys108, Trp120, Gln124, and Asp126, involved in the enzyme-substrate complex. All residues are found in the loop preceding helix H4 of **1gly**. The last three are present in every glucoamylase sequence, and Lys108 is only missing in the bacterial glucoamylase that does not show a clear homology to this region. The function of this residue might be taken by a Lys residue in the equivalent position of Tyr116 (a Phe residue in some cases) that is spacially close to Lys108 in the tridimensional model. This convergent substitution might be related with the high α-(1,6) activity of this enzyme (Ohnishi *et al.*, 1992). The modification in *S. fibuligera* equivalent to W120L led to inactivation (Itoh *et al.*, 1989). The modifications W120H, W120L, W120F, and W120Y to a decrease of the catalytic activity (Sierks *et al.*, 1989). Also the modification D126E showed a decrease of 60% in enzyme activity (Chen, 1993).

Segment S3 corresponds in the tridimensional model to the end of helix H5, a loop that contains two parts of a ß-sheet and the beginning of helix H6. The center of the loop, between the two segments forming the ß-sheet, contains the four residues, Leu177, Trp178, Glu179, and Glu180, that interact with maltose. These four residues are conserved in every glucoamylase sequence, except Leu177 which in *Clostridium* corresponds to an Arg residue. But the interaction with the subtrate by Leu177 is made by the backbone and since the *Clostridium* sequence is very divergent nothing further can be said. The modifications L177D (Bakir *et al.*, 1993) and L177H (Sierks *et al.*, 1993) reduced enzyme activity by at least 80%. Also, W178R (Sierks *et al.*, 1993) and W178D (Bakir *et al.*, 1993) showed different levels of enzyme inactivation. The mutants E179D (Bakir *et al.*, 1993) and E179Q (Sierks *et al.*, 1993) lead to complete enzyme inactivation. Mutants E180D (Bakir *et al.*, 1993) and E180Q (Sierks *et al.*, 1990) rendered substrate binding more difficult. The importance of

Glu180 in substrate binding, as described in studies with glucoamylase mutants and maltose analogues (Sierks and Svensson, 1992b), is confirmed.

Region S4 is a loop where five residues, Arg305, Tyr306, Asp309, Tyr311, and Asn315, involved in the maltose-enzyme complex can be found. This loop presents a deletion in the *Aspergillus* family of glucoamylases and an insertion in *Clostridium* glucoamylase. The five residues are conserved in all sequences, excluding Asn315, which in *Clostridium* is an Arg residue, a conservative change. Modifications Y306F and E309N (Sierks and Svensson, 1993) also led to a decrease of enzyme activity, indicating that these two residues are important in the support of the residues directly involved in binding.

The conserved segment S5 is composed of two segments involved in a ß-sheet and the beginning of helix H13 in **1gly**. The sequences of the *Saccharomyces* family present a deletion in the ß-sheet region. Four residues, Glu400, Gln 401, Lys404, and Ser411, involved in the enzyme-maltose complex, are found in this region. Glu400 is conserved in all sequences except in *Saccharomyces* where the catalytic base function may lie at position 401, a Gln residue in most of the sequences. A Glu residue in position 401 is also found in *Rhizopus* glucoamylase creating uncertainty in the position of the catalytic base in this enzyme. Lys404 is found in all the *Aspergillus* sequences being the same position occupied by an Arg residue in other sequences. Ser411 is conserved in most of the glucoamylase sequences, being a Gly residue in *S. fibuligera* and *Rhizopus*. Since this region in *Saccharomyces* shows low homology, equivalences to these sequences are uncertain.

Modifications that have been made to residues involved in the hydrogen bond net that stabilize maltose in the active site have lead to enzyme inactivation or to significant reduction in activity. Therefore, new modification of these residues sould be carefully prepared to avoid facing the same type of problems.

Critical residues in the SBD of fungal glucoamylases

The multisequence alignment of the sequences corresponding to the SBDs of fungal glucoamylases and the sequence corresponding to the C-terminal domain of *B. circulans no. 8* is shown in Figure 3. The CGTase sequence shows up to 40% identity to the glucoamylase sequences, with the N-terminal *Rhizopus* SBD sequence being the less homologous to the remaining sequences.

The C-terminal SBDs, all belonging to sequences of the *Aspergillus* family, present at least two Cys residues in conserved positions, if Cys409 of *A. niger* glucoamylase is considered as the first residue of this domain (Williamson *et al.*, 1992). As was found in the CDs of the *Aspergillus* family, where the conserved Cys residues are involved in disulfide bridges, the same phenomena can occur in the SBD where a disulfide bond has been titrated (Williamson *et al.*, 1992). Homology studies indicate that the first conserved Cys residue of the SBD of glucoamylases is found a few residues before the beginning of the C-terminal domain sequence of the CGTase, and the examination of the C-alpha structure of this domain (Schulz, 1993) shows that the second Cys residue is in an insertion in a loop close to the N-terminus so that a disulfide bond could be formed.

The six residues of the C-terminal domain of CGTases described as interacting with the substrate in the complex with maltose (Kubota *et al.*, 1990; Kubota *et al.*, 1991) are in *B. circulans* numbering Thr596, Asn601, Asn625, Gln626, Tyr631, and Trp 634. In the C-alpha structure these residues are found around a depression in the domain surface. The first residue, corresponding to Thr525 in *A. niger*, is conserved in all glucoamylase SBDs, excluding *Rizopus* where the conservative sustitute Ser is found. The second residue is the same, now Asn530 in *A. niger*, and it is also the same in three of the remaining six glucoamylase sequences. The last residue, Trp634 or Trp563 in *A. niger*, is conserved in all glucoamylase sequences. No direct equivalence is found for Tyr631, a residue present in all CGTases sequences, but assuming that the structure of this domain is the same in the two enzymes, we find a conserved Tyr residue in the SBDs — residue Tyr556 in *A. niger* — in the the same loop, probably performing the same function in substrate binding.

The function of the SBD of glucoamylase shows some similarities to the function of the C-terminal domain of CGTases. Besides the primary sequence homology, we know that the SBD of *A. niger* presents a high percentage of residues in ß-sheets (Williamson *et al.*, 1992). α- and ß-cyclodextrins, a product of CGTases, are known to inhibit the SBD ability to bind raw starch (Fukuda *et al.*, 1992). Also, deletion studies in the C-terminus of the glucoamylase and CGTase genes lead to similar types of results in the two enzymes. The deletion of the eight terminal residues in the glucoamylase SBD leads to a reduced ability to bind to raw starch (Chen *et al.*, 1992) while deletions of ten and thirteen residues in CGTase change enzyme selectivity (Kimura *et al.*, 1989). A longer deletion of 25 residues leads to the loss of ability to bind to raw starch in glucoamylase (Chen *et al.*, 1992) and a deletion of 30 terminal residues inactivates CGTase (Kaneko *et al.*, 1989). Concerning other factors affecting SBD activity, chemical modification of Trp590 and Trp615 reduces starch-binding activity (Svensson *et al.*, 1986). These two Trp residues are also found in all CGTase sequences. Other subsites in the SBD of fungal glucoamylases, able to singularly bind a sugar residue, might then be located close to the C-terminus of the enzyme in the region affected by the deletions.

Linker region in the Aspergillus family of glucoamylases

The multisequence alignment corresponding to the linker region in glucoamylases of the *Aspergillus* family is shown in Figure 4. The C-terminus of the CD is estimated as being Pro467 (in *A. niger* numbering), the last conserved residue in this domain before a region of great variability. Linker regions may present great variability (Zalkin *et al.*, 1984) since the connective function is not dependent on a specific sequence. This situation agrees with the results of deletion studies where the excision of the segment Ala462 to the C-terminus lead to destabilization of the enzyme (Evans *et al.*, 1990). The structure of the O-glycosylation region after Ala471 has been studied (Williamson *et al.*, 1992), indicating that this region is extended and rigified bythe O-glycosylation sugars.

O-glycosylation is found before Pro467 and after Cys509, extending into both the CD and the SBD. The number of Ser and Thr residues in *Humicola grisea var. thermoidea* and *Hormoconis resinae* is smaller than in the four *Aspergillus* sequences, an indication of the lesser role of O-glycosylation in these structures. These two sequences present a shorter linker region together with the *A. oryzae* sequence, and this fact indicates that the long and conserved Ser/Thr-rich region of the remaining *Aspergillus* sequences might be recent in evolutionary terms. The deletion of part of the O-glycosylated region in *A. niger* glucoamylase caused problems of secretion when expressed in yeast without affecting enzyme activity (Libby *et al.*, 1993), linking O-glycosylation to secretion.

The length of the linker, whose structure varies significantly in the shorter sequences, is five to fourty residues. Assuming that this structure stays extended in solution, for a mean length of 2.73 Å per linker residue (Argos, 1990), the length of the linker will vary between 14 and 110 Å. Knowing from crystallographic data that the CD can be presented as a globular structure with a 65 Å diameter and that the longer length of the C-terminal domain is approximately 35 Å, the variations in length of the linker indicate the two domains composing the fungal glucoamylases are functionally independent.

ACKNOWLEDGEMENTS

P.M.C. was supported by grant B.D./1057/90-IF of Programa CIENCIA, Junta Nacional de Investigação Científica e Tecnológica, Lisboa, Portugal.

REFERENCES

Aleshin A., Golubev A., Firsov L.M. and Honzatko R.B. (1992) *J. Biol. Chem.*, 267, 19291-19298.
Aleshin A. (1992) personal communication.
Argos, P. (1990) *J. Mol. Biol.*, 211, 943-958.

Ashikari T., Nakamura N., Tanaka Y., Kiuchi N., Shibano Y., Tanaka T., Amachi T. and Yoshizumi H. (1986) *Agric. Biol. Chem.*, 50, 957-964.
Bakir U., Coutinho P.M., Sullivan P.A., Ford C. and Reilly P.J. (1993) *Protein Eng.*, in press.
Bégin P., Millet J., Chauvaux S., Salamitou S., Tokalides K., Navas J., Fujino T., Lemaire M., Raynaud O., Daniel M.K. and Aubert J.-P. (1992) *Biochem. Soc. Transac.*, 20, 42-46.
Berka R.M., Rey M.W., Thompson S.A., Gray G.L., Carmona C.L., Power S.D. (1993) unpublished.
Boël E., Hjort I., Svensson B., Norris F., Norris K.E. and Fiil N.P. (1984) *EMBO J.*, 3, 1097-1102.
Bock K. and Pedersen H. (1987) *Acta Chem. Scandinavica*, B41, 617-628.
Chen L.-J. (1992) personal communication.
Chen H.-M.(1993) personal communication.
Chen H.-M., Bakir U., Oates P.D., Reilly P.J. and Ford C. (1993) *Biotechnol. Bioeng.*, accepted for publication.
Dalmia B.K. and Nikolov Z.L. (1991) *Enzyme Microb. Technol.*, 13, 982-990.
Devereux J., Haeberli P. and Smithies O. (1984) *Nucl. Acids Res.*, 12, 387-395.
Dohmen R. J., Strasser A.W.M., Dahlems U.M. and Hollenberg C.P. (1990) *Gene*, 95, 111-121.
Evans R., Ford C., Sierks M., Nikolov Z. and Svensson B. (1990) *Gene*, 91, 131-134.
Feng D.D. and Doolittle R.F. (1987) *J. Mol. Evol.*, 25, 351-360.
Fukuda K., Teramoto Y., Goto M., Sakamoto J., Mitsuiki S. and Hayashida S. (1992) *Biosci. Biotech. Biochem.*, 56, 556-559.
Gašperík J. and Hostinová E. (1993) *Curr. Microbiol.*, in press.
Genetics Computer Group (1991) *Program Manual for the GCG Package - Version 7 , April 1991*, 575 Science Drive - Suite B, Madison, Wisconsin 53711.
Harris E.M.S., Aleshin A.E., Firsov L.M. and Honzatko R.B. (1993) *Biochemistry*, 32, 1618-1626.
Hata Y., Tsuchiya K., Kitamoto K., Gomi K., Kumagai C., Tamura G. and Hara S. (1991) *Gene*, 108, 145-150.
Hayashida S., Kuroda K., Ohta K., Kuhuara S., Fukuda K. and Sakaki Y. (1989) *Agric. Biol. Chem.*, 54, 923-929.
Henrissat B. and Mornon J.-P. (1990) *Trichoderma reesei Cellulases: Biochemistry, Genetics, Physiology and Application*, eds. Kubicek C.P., Eveleigh D.E., Esterbauer H., Steiner W. and Kubicek-Pranz E.M. , Royal Society of Chemistry, 12-29.
Hostinová E., Balanová J. and Gasperíc J. (1991) *FEMS Microb. Lett.*, 83, 103-108.
Itoh T., Ohtsuki I., Yamashita I. and Fukui S. (1987) *J. Bacteriol.*, 169, 4171-4176.
Itoh T., Sakata Y., Akada R., Nimi O. and Yamashita I. (1989) *Agric. Biol. Chem.*, 53, 3159-3167.
Jutsjoki V.V. and Torkkeli T.K. (1992) *FEMS Microbiol. Lett.*, 99, 237-244.
Juy M., Amit A.G., Alzari P.M., Poljack R.J., Claeyssens M., Bégin P. and Aubert J.-P. (1992) *Nature*, 357, 89-91.
Kaneko T., Song K.B., Hamamoto T., Kudo T. and Horikoshi K. (1989) *J. Gen. Microbiol.*, 135, 3447-3457.
Kimura K., Kataoka S., Nakamura A., Takano T., Kobayashi S. and Yamane K. (1989) *Biochem. Biophys. Res. Commun.*, 161, 1273-1279.
Klein C. and Schulz G. E. (1991) *J. Mol. Biol.*, 217, 737-750.
Klein C., Hollander J., Bender H. and Schulz G.E. (1992) *Biochemistry*, 31, 8740-8746.
Kubota M., Matsuura Y., Sakai S. and Katsube Y. (1990) Abstract, *Int. Symp. Cereal & Other Plant Carbohydrates*, Kagoshima, 132-133.
Kubota M., Matsuura Y., Sakai S. and Katsube Y. (1991) *Denpun Kagaku*, 38, 141-146.
Lemieux R.U. (1992) personal communication.
Libby C., Geurkink C., Ford C. and Reilly P.J. (1993) in preparation.
Matsuura Y., Kusunoki M., Harada W. and Kakudo M. (1984) *J. Biochem.*, 95, 697-702.
Naim H.Y., Nierman T., Kleinhans U., Hollenberg C.P. and Strasser A.W.M. (1991) *FEBS Lett.*, 294, 109-112.
Nitschke L., Heeger K., Bender H. and Schulz G.E. (1990). *Appl. Microbiol. Technol.*, 33, 542-546.
Nunberg J.H., Meade J.H., Cole G., Lawyer F.C., McCabe P., Schweickart V., Tal R., Wittman V.P., Flatgaard J.E. and Innis A. (1984) *Mol. Cell. Biol.*, 84, 2306-2315.
Ohnishi H., Sakai H. and Ohta T. (1991) *Agric. Biol. Chem.*, 55, 1901-1902.
Ohnishi H., Kitamura H., Minowa T., Sakai H. and Ohta T. (1992) *Eur. J. Biochem.*, 207, 413-418.

Pazur J. H. and Ando T. (1959) *J. Biol. Chem.* 234, 1966-1970.
Petukov M.G. and Mazur A.K. (1992) *Molekularnaya Biologiya*, 26, 292-299.
Raimbaud E., Buléon A., Pérez S. and Henrissat B. (1989) *Int. J. Biol. Macromol.*, 11, 217-225.
Sakagushi K., Takagi M., Horiuchi H. and Gomi K. (1992) *Applied Molecular Genetics of Filamentous Fungi*, eds. Kinghorn J.R. and Turner G. (Blackie, Glasgow), 54-99.
Shibuya I., Gomi K., Iimura Y., Takahashi K., Tamura G. and Hara S. (1990) *Agric. Biol. Chem.*, 54, 1905-1914.
Schulz G.E. (1993) personal communication.
Sierks M.R., Ford C., Reilly P.J. and Svensson B. (1989) *Protein Eng.*, 2, 621-625.
Sierks M.R., Ford C., Reilly P.J. and Svensson B. (1990) *Protein Eng.*, 3, 193-198.
Sierks M.R., Ford C., Reilly P.J. and Svensson B. (1993) *Protein Eng.*, 6, 75-84.
Sierks M.R. and Svensson B. (1992a) US Patent 5,162,210.
Sierks M.R. and Svensson B. (1992b) *Protein Eng.*, 5, 185-188.
Sierks M.R. and Svensson B. (1993) *Biochemistry*, 32, 1113-1117.
Svensson B., Clarke A.J. and Svendsen I. (1986) *Carlsberg Res. Commun.*, 51, 61-73.
Svensson B., Jespersen H., Sierks M.R. and MacGregor E.A. (1989) *Biochem. J.*, 264, 309-311.
Tanaka Y., Ashikari T., Nakamura N., Kiuchi N., Shibano Y., Amachi T. and Yoshizumi H. (1986) *Agric. Biol. Chem.*, 50, 965-969.
Wiliamson G., Belshaw N.J. and Williamson P. (1992) *Biochem. J.*, 282, 423-428.
Wood T. M. (1992) *Biochem. Soc. Trans.*, 20, 46-53.
Yamashita I., Suzuki K. and Fukui S. (1985) *J. Bacteriol.*, 161, 567-573.
Yamashita I., Nakamura M. and Fukui S. (1987) *J. Bacteriol.*, 169, 2142-2149.
Yamashita I. (1989) *Agric. Biol. Chem.*, 53, 483-489.
Zalkin H., Paluh J.L., van Cleemput M., Moye W.S. and Yanofsky C. (1984) *J. Biol. Chem.*, 259, 3985-3992.

FIGURES AND TABLES

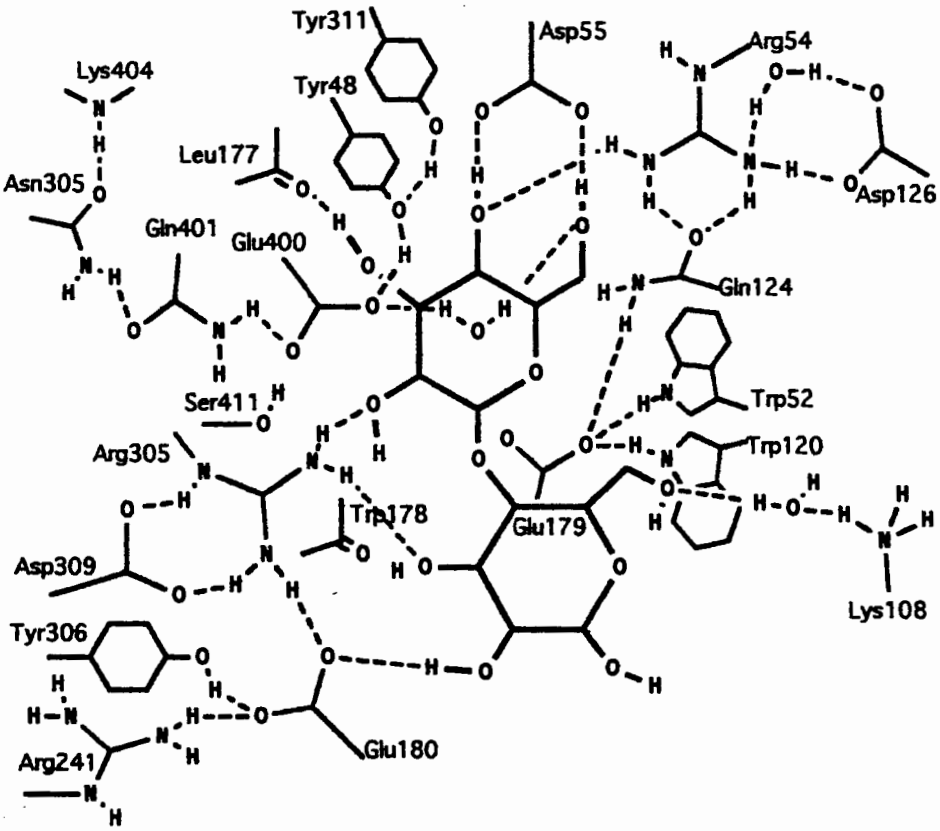


Figure 1 - Hydrogen bonds in the glucoamylase-maltose complex.

```
      Region S1
                                                          *   * **
1gly    26 - GAWVS.............GADSGIVV.ASPSTDNPD.........YFYTWTRDSGLV - 59
1gly.ss 26 -                      ßßßß                      αααααααα - 59
aspak   26 - GAWVS.............GADSGIVV.ASPSTDNPD.........YFYTWTRDSGLV - 59
aspsh   26 - GAWVS.............GADSGIVV.ASPSTDNPD.........YFYTWTRDSGIV - 59
aspng   26 - GAWVS.............GADSGIVV.ASPSTDNPD.........YFYTWTRDSGLV - 59
aspor   27 - GQSAQ.............GASPGIVVIASPSKSDPD.........YFYTWTRDSGLV - 61
humgr   26 - GKAAP.............GAAAGVVI.ASPSRTDPP.........YFFTWTPDAALV - 59
hrrsp   25 - GSAVP.............GAGAGFVV.ASPSKANPD.........YFYTWSRDSALT - 58
sacfi   41 - EGQFN.............NGVPGTVI.ASPSTSNPD.........YYYQWTRDSAIT - 74
sacfk   41 - EGQFN.............DGVPGTVI.ASPSTSNPD.........YYYQWTRDSAIT - 74
sacce  103 - SAVYP.............SISPGVVI.ASPSQTHPD.........YFYQWIRDSALT - 136
sacdi  339 - SAVYP.............SILPGVVI.ASPSQTHPD.........YFYQWIRDSALT - 372
rhior  168 - GSA.................TGFIA.ASLSTAGPD.........YYYAWTRDAALT - 196
clost  290 - ANSLYYNSMMILKASEDKTNKGAYI.ASLSIPWGDGQGDDNTGGYHLVWSRDLYHV - 344

      Region S2                                          Region S3
               *           *   * *                                ****
1gly   105 - GEPKFNVDETAYTGSWGRPQRDGPALRATA - 134    162 - DLSYVAQYWNQTGYDLWEEVNGSSFFTIAVQHRAL - 196
1gly.ss105 -                      αααααααααα - 134    162 - αααααααα    ßßßß    ßßßßßαααααααααααα - 196
aspak  104 - GEPKFNVDETAYTGSWGRPQRDGPALRATA - 133    161 - DLSYVAQYWNQTGYDLWEEVNGSSFFTIAVQHRAL - 195
aspsh  104 - GEPKFNVDETAYAGSWGRPQRDGPALRATA - 133    161 - DLSYVAQYWNQTGYDLWEEVNGSSFFTIAVQHRAL - 195
aspng  105 - GEPKFNVDETAYTGSWGRPQRDGPALRATA - 134    162 - DLSYVAQYWNQTGYDLWEEVNGSSFFTIAVQHRAL - 196
aspor  106 - GEPKFNVDETAFTGAWGRPQRDGPALRATA - 135    163 - DLSYVAQYWSQSGFDLWEEVQGTSFYTVAVSHRAL - 197
humgr   92 - GEAKFNVDLTAFTGEWGRPQRDGPPLRAIA - 121    149 - DLAYTAQYWNETGFDLWEEVPGSSFFTIASSHRAL - 183
hrrsp  105 - GEPKFMVDGTRFNGPWGRPQRDGPALRAIA - 134    162 - DLSYVGQYWNQSGFDLWEETYASSFFTIQNQHRAL - 196
sacfi  124 - GEPKFNTDGSAYTGAWGRPQNDGPALRAYA - 153    193 - DLEYVIGYWDSTGFDLWEENQGRHFFTSLVQQKAL - 227
sacfk  124 - GEPKFNTDGSAYTGAWGRPQNDGPALRAYA - 153    193 - DLEYVIGYWDSTGFDLWEENQGRHFFTSLVQQKAL - 227
sacce  183 - GDPKWNVDNTAFTEDWGRPQNDGPALRSIA - 212    247 - DLRFIIDHWNSSGFDLWEEVNGMHFFTLLVQLSAV - 281
sacdi  419 - GDPKWNVDNTAFTEPWGRPQNDGPALRSIA - 448    483 - DLRFIIDHWNSSGFDLWEEVNGMHFFTLLVQLSAV - 517
rhior  249 - GEPKFNPDGSGYTGAWGRPQNDGPAERATT - 268    297 - DLDYVVNVWSNGCFDLWEEVNGVHFYTLMVMRKGL - 331
clost  377 - .....NTWISGK.PYWTGIQLDEQAD.... - 396    427 - LADFIIKMGPKTGQERWEEIGGYSPATMAAEVAGL - 451

      Region S4                                          Region S5
                 **  *                 *                          **  *       *
1gly   300 - AVAVGRYPEDSYYN.............GNPWFQS - 320    396 - GSMSEQFDKSDGDEL.SARDLTWSYAALLTA - 425
1gly.ss300 -                              ααα - 320    396 - ßßßßßßßßß  ßßßßß ßßßßßαααααααααααα - 425
aspak  299 - AVAVGRYPEDSYYN.............GNPWFQS - 319    395 - GSLSEQFDKSDGDEL.SARDLTWSYAALLTA - 424
aspsh  299 - AVAVGRYPEDSYYN.............GNPWFLC - 319    395 - GSLSEQFDKSDGDEL.SARDLTWSYAALLTA - 424
aspng  300 - AVAVGRYPEDTYYN.............GNPWFLC - 320    396 - GSMSEQYDKSDGEQL.SARDLTWSYAALLTA - 425
aspor  301 - AVAVGRYPEDSYYN.............GNPWFLT - 321    397 - GSMAEQYTKTDGSQT.SARDLTWSYAALLTA - 426
humgr  291 - AVAVGRYSEDVYYN.............GNPWYLA - 311    387 - GALAEQYDRNTGKPD.SAADLTWSYSAFLSA - 416
hrrsp  301 - GVAVGRYAEDVYMG.............GNPWYLI - 321    399 - GSLSEQFNRDTGTPL.SAIDLTWSYAAFITM - 428
sacfi  340 - GAAIGRYPEDVYNGDGSSE........GNPWFLA - 365    452 - GSLNEQLNRYTGYST.GAYSLTWSSGALLEA - 481
sacfk  340 - GAAIGRYPEDVYNGDGSSE........GNPWFLA - 365    452 - GSLNEQLNRYTGYST.SAYSLTWSSGALLEA - 481
sacce  396 - GIALGRYPEDVYDGYGFGE........GNPWVLA - 421    472 - GYLILEFN......T.PAFNQT.IQKIFQLA - 494
sacdi  632 - GIALGRYPEDVYDGYGVGE........GNPWVLA - 657    708 - GYLILEFN......T.PAFNQT.IQKIFQLA - 730
rhior  448 - GNSIGRYPEDTYNGNGNSQ........GNSWFLA - 463    540 - GSLAEEFDRTTGLST.GARDLTWSHASLITA - 569
clost  566 - GPSWYRYNHDGYGEPSKTELYHGAGKGRLWPLL - 598    628 - GIISEQVWEDTGLPTDSASPLNWAHAEYVVL - 658
```

Figure 2 - Conserved regions S1 to S5 in the catalytic domain of glucoamylases. 1gly - *Aspergillus awamori var. X100* (1gly.ss - secondary structure); aspak - *Aspergillus awamori var. kawachi*; aspsh - *Aspergillus shirousami*; aspng - *Aspergillus niger /Aspergillus awamori*; aspor - *Aspergillus oryzae*; humgr - *Humicola grisea var. thermoidea*; hrrsp - *Hormoconis resinae*; sacfi - *Saccharamycopsis fibuligera* ; sacfk - *Saccharomycopsis fibuligera KZ*; sacce - *Saccharomyces cerevisae* ; sacdi - *Saccharomyces diastaticus*; rhior - *Rhizopus oryzae*; clost - *Clostridium sp.* .
* - Residues of the active site involved in maltose binding.

```
                                    *   *                                                  **    *
baccr          581 - ...GDQVTVRFVVNN.ASTTLGQNLYL........TGNVAELG..NWSTGSTAIG..PAFNQVIHQYPT - 633
baccr.ss       581 -    ßßßßßßßßßß          ßßßß         ßßß                  ßß  ßß           - 633
aspng          509 - CTTPTAVAVTFDL.T.ATTTYGENIYL........VGSISQLG..DWET.SDGIA..LSADKYTSSDPL - 562
aspak          508 - CTTPTAVAVTFDL.T.ATTTYGENIYL........VGSISQLG..DWET.SDGIA..LSADKYTSSNPL - 561
aspsh          508 - CTTPTAVAVTFDL.T.ATTTYGENIYL........VGSISQLG..DWET.SDGIA..LSADKYTSSNPP - 561
aspor          480 - CQVPTTVSVTFAV.K.ATTVYGESIKI........VGSISQLG..SWNP.SSATA..LNADSYTTDNPL - 533
humgr          479 - CADASEVYVTFNE.R.VSTAWGETIKV........VGNVPALG..NWDT.SKAVT..LSASGYKSNDPL - 532
hrrsp          476 - ....CQVSITFNI.N.ATTYYGENLYV........IGNSSDLG..AWNI.ADAYP..LSASAYTQDRPL - 532
rhior            1 - ASIPSSASVQLDSYNYDGSTFSGKIYVKNIAYSKKVTVIYADGSDNWNNNGNTIAASYSAPISGSNYEY - 69

                     *
baccr          634 - WYYDVSVPAGK.QLEFKFFKKNGS..TITWESGSNHTFTTP......ASGTATVTVNWQ..........  - 684
baccr.ss       634 - ßßßßßßßßß ß ßßßßßßßßß        ßßß    ßßßßß           ßßßßßßß            - 684
aspng          563 - WYVTVTLPAGES.FEYKFIRIE.SDDSVEWESDPNREYTVPQAC..GTSTA.TVTDTWR..........  - 616
aspak          562 - WYVTVTLPAGES.FEYKFIRVE.SDDSVEWESDPNREYTVPQAC..GESTA.TVTDTWR..........  - 615
aspsh          562 - WYVTVTLPAGES.FEYKFIRVE.SDDSVEWESDPNREYTVPQAC..GESTA.TVTDTWR..........  - 615
aspor          534 - WTGTINLPAGES.FQYKFIRVQ..NGAVTWESDPNRKYTVPSTC..GVKSA.VQSDTWR..........  - 586
humgr          533 - WSITVPIKATGSAVQYKYIKVG.TNGKITWESDPNRSITLQTASSAGKCAAQTVNDSWR..........  - 590
hrrsp          533 - WSAAIPLNAGE.VISYQYVRQEDCDQPYIYET.VNRTLTVP.AC...GGAAVTTDDAWMGPVGSSGNC  - 587
rhior           70 - W....TFSASINGIKEFYIKYEVSG.....................................------  - 90
```

Figure 3 - Starch-binding domain of fungal glucoamylases compared to the C-terminal domain of *Bacillus circulans no.8* CGTase. CGTase: **baccr** - *Bacillus circulans* no. 8 (**baccr.ss** - secondary structure). Other abbreviations as in Figure 2. * - Residues involved in maltose binding in CGTase.

```
                  **        ** * * * ** * *
1gly     438 - GETSASS.VPGTCAATSASGTYSSVTVTSWPSIVA..................................... - 471

                 * **        **      ** * *   *   *  **** * * * * **** ** * * ********* **
aspng    438 - GETSASS.VPGTCAATSAIGTYSSVTVTSWPSIVATGGTTTTATPTGSGSVTSTSKTTATASKTSTSTSSTSCTTPTAVA - 516
aspak    437 - GETSASS.WPGTCAATSASGTYSSVTVTSWPSIVATGGTTTTATTTGSGGVTSTSKTTTTASKTSTTTSSTSCTTPTAVA - 515
aspsh    437 - GETSASS.VPGTCAATSASGTYSSVTVTSWPSIVATGGTTTTATTTGSGGVTSTSKTTTTASKTSTTTSSTSCTTPTAVA - 515
aspor    439 - GETAATS.IPSACSTTSASGTYSSV.VTSWPTISGYPG.....................APDSPCQVPTTVS - 487
humgr    429 - RASVAKSQLPSTCSRIEVAGTYVAATSTSFPS.....................KQTPNPSAAPSPSPYPTACADASEVY - 486
hrrsp    441 - GSRNALPP.PTTCSASSTPGIYTPATAAGAPNV.....................TSSCQ....VS - 479
```

Figure 4 - O-glycosylation region in glucoamylases of the *Aspergillus* family. Abbreviations as in Figure 2. ▬ - Linker Region ; * - O-glycosylated residues.

Table 1 - Amino acid residue identity between fungal and yeast glucoamylase sequences after multisequence progressive alignment. The legend is: **1gly** - *Aspergillus awamori var. X100* ; **aspsh** - *Aspergillus shirousami*; **aspak** - *Aspergillus awamori var. kawachi*; **aspng** - *Aspergillus niger /Aspergillus awamori*; **aspor** - *Aspergillus oryzae*; **humgr** - *Humicola grisea var. thermoidea*; **hrrsp** - *Hormoconis resinae*; **rhior** - *Rhizopus oryzae*; **sacfl** - *Saccharamycopsis fibuligera* ; **sacfk** - *Saccharomycopsis fibuligera KZ*; **sacce** -*Saccharomyces cerevisae* ; **sacdi** -*Saccharomyces diastaticus.* Fungal families are indicated in gray background and yeast families in open boxes.

	1gly	aspsh	aspak	aspng	aspor	humgr	hrrsp	sacfi	sacfk	sacce	sacdi	rhior
1gly	470	465	467	454	375	282	268	178	179	152	153	162
aspsh		615	606	593	443	337	318	180	181	150	151	161
aspak			615	590	442	337	317	178	179	152	153	162
aspng				616	441	341	324	183	184	155	156	166
aspor					586	329	309	176	177	146	146	172
humgr						590	305	163	164	139	138	167
hrrsp							587	176	177	141	143	167
sacfi								492	485	203	205	197
sacfk									492	204	206	196
sacce										510	473	160
sacdi											746	171
rhior												579

CELLULAR RESPONSES OF INSECT CELL *SPODOPTERA FRUGIPERDA* TO ENVIRONMENTAL STRESSES

Paul Yeh, Grace Y. Sun[1], Gary A. Weisman[1], Rakesh Bajpai

Department of Chemical Engineering
[1]Department of Biochemistry
University of Missouri, Columbia, Missouri

INTRODUCTION

The insect cell line *Spodoptera frugiperda* (Sf9) can be used for the production of viral insecticides and recombinant proteins via baculovirus transfer vector system (1). The sensitivity of Sf9 cells to environmental stresses has been investigated using Couette viscometer and bubble column (2), although information on cell survival and metabolism after stress is limited. In many instances, it is necessary to study cells that are grown in a serum-free medium supplemented with pluronic F68, a non-ionic surfactant which is used as an emulsifier. The detergent can protect these insect cells from detrimental effects of bubbling and vortexing especially when the cells are cultured in large reaction flasks. This study is to examine the effects of F68 as well as well-defined shear stress conditions on cell membrane fatty acid composition and lipids metabolism.

EXPERIMENTAL PROCEDURE

CELL CULTURE

Sf9 cells were grown in stationary T175 culture flasks containing TNM-FH medium supplemented with heat inactivated 10% fetal bovine serum (FBS). Cells were grown at 28°C and were subcultured twice a week (3).

LIPID ANALYSIS

After subculturing in 150 mm dishes for 4 h, cells were washed with 2 ml of Hank-balanced salt solution (HBS) three times and then scraped into a borosilicate glass tube. Eight ml of chloroform-methanol (2:1, v/v) was added followed by vortexing the cell sample. The lower organic phase was collected and filtered through an anhydrous Na_2SO_4 column. The neutral lipids in the extract were separated on silica gel HPTLC plates (Whatman HP-K) with a solvent system comprising of hexane-ethyl ether-acetic acid (85:15:2, by vol). To further identify individual phospholipid components, two dimensional HPTLC was applied using first a solvent system comprising of chloroform-methanol-acetone-ammonium hydroxide (75:40:10:10, by vol) and in a second solvent system comprising of chloroform-methanol-acetone-acetic acid:ammonium acetate (70:30:27.5:2.3:5, by vol) (4). The plates were air-dried and lipids were visualized by exposure of HPTLC plates to 2'7'dichlorofluorescein and viewing the spots under UV light. Fatty acid of individual lipids were converted to methyl esters by base-methanolysis (0.2N NaOH-methanol) and subsequently analyzed by gas chromatography. Fatty acids methyl esters were quantified by using the 17:0 fatty acyl methyl ester as the internal standard. FAME were analyzed using HP 5890 gas chromatograph equipped with a fame ionization detector and a SP 2330 fuse silica capillary column.

UPTAKE OF [^{14}C] OLEIC ACID BY SF9 CELLS

Besides analysis of the phospholipid and acyl group composition, the effects of detergent and shear stress on cell membrane lipids were evaluated by incubating cells in the presence [1-^{14}C]oleic acid (0.1 μCi/10^6 cells) in growth medium for 1 h at 28°C. After incubation, lipids were extracted and separated by HPTLC as described above. Lipid spots were removed from the HPTLC plate for counting of radioactivity using Bechman LS 5801 scintillation spectrometer.

SHEAR APPLICATION

(1) Chemical stress: Sf9 cells were grown in a complete culture medium containing 0.2% F68. After incubation for 4 days, the cells were washed twice by HBS and lipids were extracted and separated by chloroform-methanol as described above. (2) Mechanical stress: Sf9 cells in culture were harvested by light trypsinization and were loaded into a cocylindrical Rheoset viscometer. The shear stress in the compartment was determined by the shear rate of the inner cylinder. Control samples were loaded similarly into the viscometer without applying shear

stress. After stress, cells were removed from the viscometer and transferred to test tubes. They were washed and subsequently incubated in a medium containing 1 ml of culture medium with 0.1 μCi [^{14}C] oleic acid for 1 h. Control and sheared cells were washed and processed as described above.

RESULTS AND DISCUSSION

Fatty acid composition in phospholipids of SF9 cells

SF9 cells were characterized by phospholipids containing a high proportion (over 70%) of monounsaturated fatty acids such as 16:1 and 18:1. These monounsaturated fatty acids had been identified using $AgNO_3$ argentation HPTLC. Interestingly, polyunsaturated fatty acids were found to be less than 2% of total phospholipids in SF9 cells (Fig. 1). This suggests arachidonic acid (20:4) metabolism may not play an important role in insect cells.

Effect of mechanical shear stress on incorporation of ^{14}C-oleic acid into phospholipids and triacylglycerol

Radiolabeled oleic acid (0.1 μCi/ml) was actively taken into all phospholipid species and the labeling of phospholipids reached a peak at 5 hr. The labeling of TG reached a peak at 2 hr and then declined to a steady level at 10 hr (Fig. 2a). There was an increase uptake of labeled oleic acid by SF9 membrane lipids, especially after 40 min of applied shear stress (Fig. 2b). Addition of pluronic F68 (0.2%) into the medium also showed an increase in taking up more labeled oleic acid in the phospholipids than control (Fig. 2c). The results are consistent with the hypothesis that shear stress at sublethal level as well as surfactant F68 may facilitates the lipid uptake into the membrane which attributes to cell growth.

Effect of shear stress on uptake of ^{14}C-oleic acid into membrane lipids

Besides the increase in radioactivity, applied shear stress resulted in an increase in labeling of PE and an decrease in labeling of PC (Table 1). Mechanical shear stress stimulates a diacylglycerol-independent pathway of protein kinase C activation in platelet. The shifts of labeled oleic acid into different lipid components suggests shear stress may alter lipid metabolic pathway. It was found the cell exposed to 10% FBS had higher PC and lower labeling of PE than those grown in 2.5% serum (Fig. 3). SF9 cells exposed to F68 did not show any difference in the distribution of the labeling in phospholipid species.

Effect of surfactant F68 on phospholipids and fatty acid composition of SF9 cells

Exposure of SF9 cells to 0.2% of F68 resulted in an increase in the level of PC and a decrease in levels of PS/PI (Fig. 4a). Although PE did not show obvious change in levels, there was a decrease in proportion of 16:1 with concomitant increase in 18:1 due to exposure to F68 (Fig. 4b). Similar changes in fatty acid composition were observed in PS/PI (Fig. 4c). It was noted that despite the increase in mass, there was no change in fatty acid composition of PC.

CONCLUSIONS

1. Analysis of Sf9 cell membrane lipids revealed high proportions of monoenoic fatty acids (16:1, 18:1) in the phospholipids. These cells are practically deprived of polyunsaturated fatty acids. This lipid profile seems to be different from that of most mammalian cells.

2. There is a time-dependent uptake of ^{14}C-labeled oleic acid into phospholipids and triacylglycerols of Sf9 cells. Addition of F68 (0.2% w/v) in the presence of serum (2%) only facilitated the uptake of labeled oleic acid into phospholipids but also altered the phospholipid fatty acid composition.

3. After shear Sf9 cells of 4 dynes/cm^2 for 40 min, incorporation of oleate into membrane lipids was increased. Among the phospholipids, uptake of label into phosphatidylethanolamine was increased.

4. It is concluded that subtle changes in membrane lipids and fatty acids occur in Sf9 cells due to environmental stress. Further studies will be needed to relate these changes to their ability to survive these insults.

REFERENCES

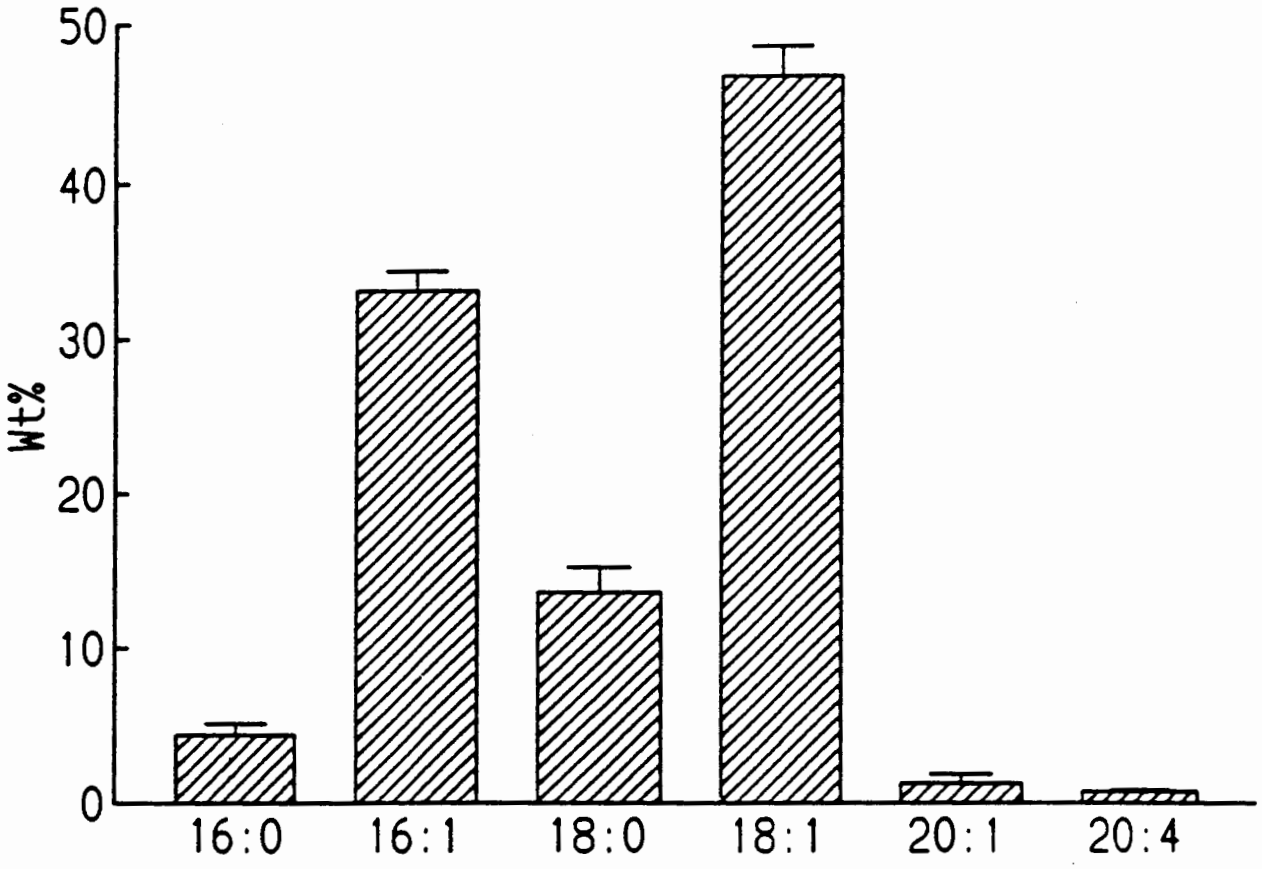
1. Maiorella, B., Inlow, D., Shauger, A., Harano, D. (1988) Bio/Technol. 6:1406-1412.
2. Murhammer, D. W., Goochee, C. F. (1991) Biotechnol. Progr. 7:205-211.
3. Luckow, V.A., Summer, M.D. (1988) Bio/Technol. 6:47-54.
4. Sun, G. Y. (1988) From: Lipids and Related Compounds, Ed. by Boulton, A.A., Baker, G. B., and Horrocks, L.O. The Humana Press, pp 63-82.

TABLE 1

Effect of shear stress on incorporation of ^{14}C oleic acid into membrane lipids

Conditions	TG	PS	PI	PC	PEpl	PE
			%			
Cntl	12.8 ± 1.6	2.1 ± 0.1	5.3 ± 0.2	52.6 ± 1.2	0.6 ± 0.2	26.7 ± 0.4
S20	9.2 ± 1.7	2.5 ± 0.2	6.6 ± 1.0	49.7 ± 1.0	0.7 ± 0.1	31.3 ± 0.9
Cnt2	5.9 ± 0.7	2.9 ± 0.2	8.6 ± 0.4	48.6 ± 1.7	1.1 ± 0.6	32.9 ± 1.4
S40	6.8 ± 0.2	2.6 ± 0.1	6.0 ± 0.1	45.8 ± 0.6	0.7 ± 0.4	38.1 ± 0.6

Fig. 1 Fatty acid composition in phospholipids of Sf9 cells



Sf9 cells (1×10^7) were grown in TNM-FH medium containing 10% fetal bovine serum (FBS). After 4 days of incubation, cells were harvested by trypsin and resuspended in 1 ml HBS. Cellular lipids were extracted by adding 4 ml of chloroform/methanol (2:1, v/v). Cellular lipid levels were analyzed as described in "Experiment Procedures". The data are expressed as percent distribution of the fatty acids components in total phospholipids. Results are mean ± SD of triplicate determinations.

Remark: Sf9 cells are characterized by phospholipids containing a high proportion (over 70%) of monounsaturated fatty acids such as 16:1 and 18:1. These monounsaturated fatty acids have been identified using $AgNO_3$ argentation HPTLC.

Fig. 2a Incorporation of ^{14}C-oleic acid into phospholipid and triacylglycerol

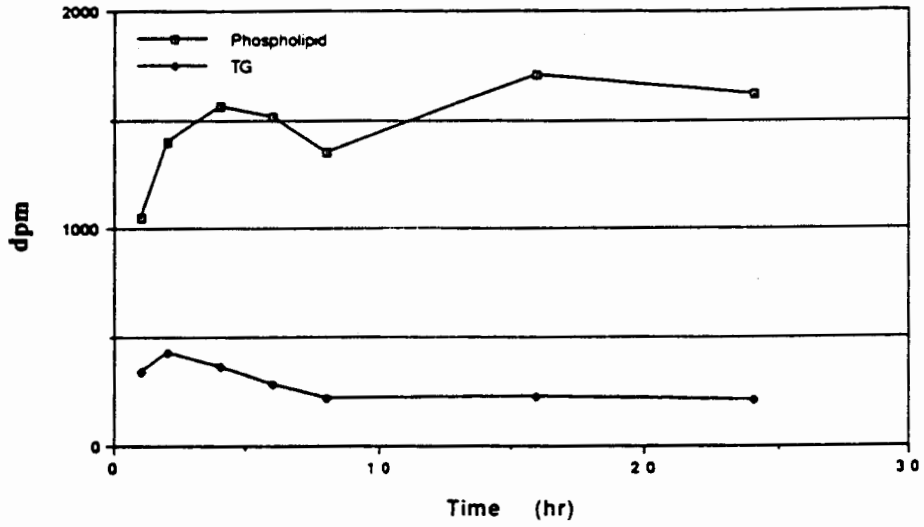
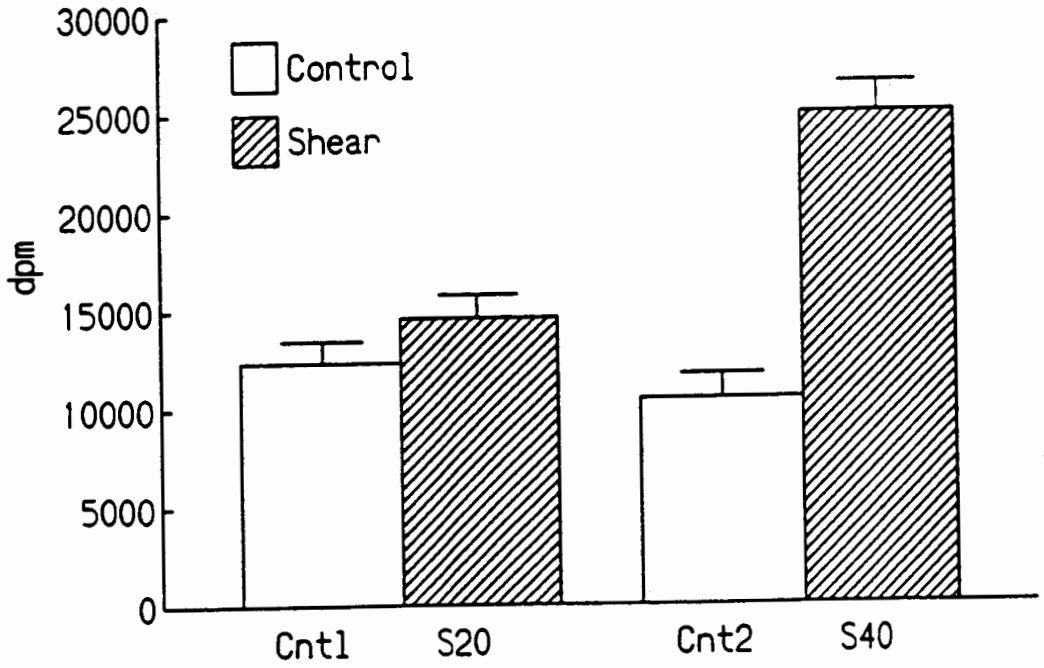
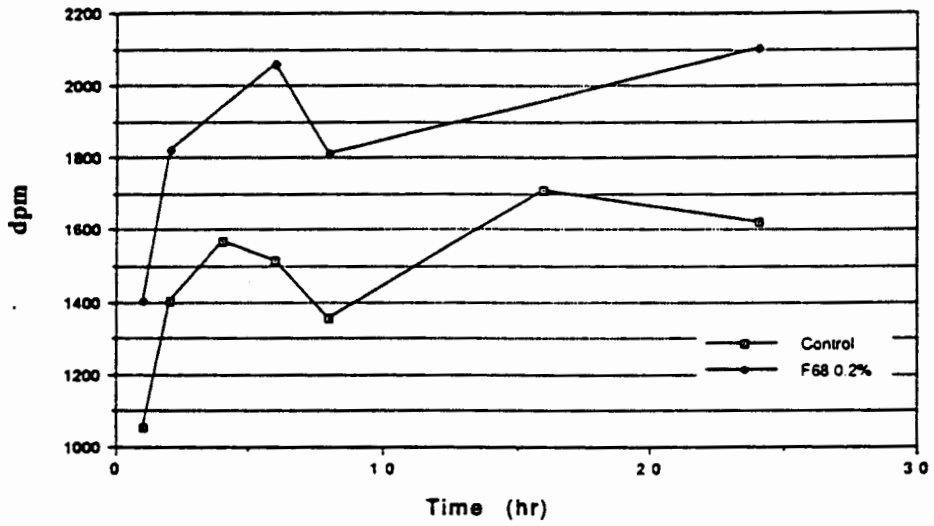


Fig. 2b Uptake of ^{14}C-oleic acid by Sf9 membrane lipids



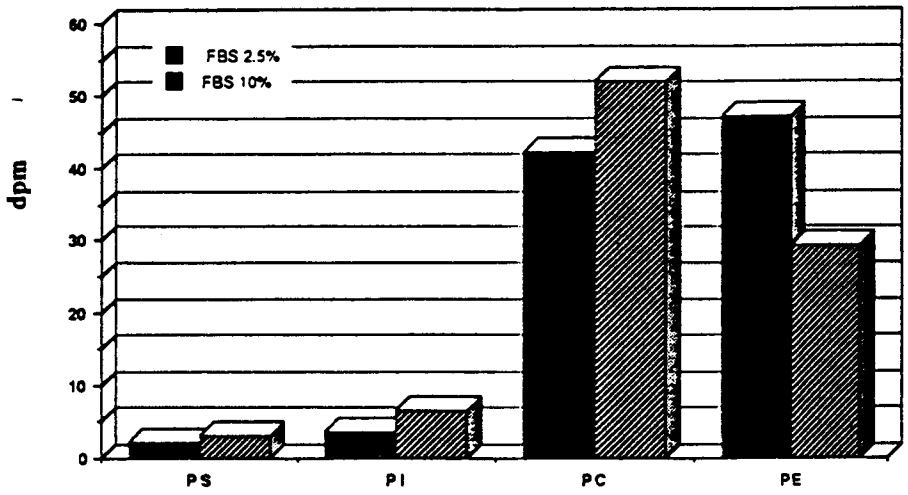
After four days incubation in complete medium, Sf9 cells were harvested by trypsinization and loaded into a viscometer under given shear stresses of 5 dynes/cm^2 for 20 and 40 min. Control cells (Cntl) are also loaded without applying shear. After applying shear, cells (1×10^6) were transferred into test tubes and incubated with [^{14}C]oleic acid (0.1 μCi) as described in "Experimental Procedures". After labeling, cells were washed twice, resuspended in 1 ml HBS, and the lipids were extracted and analyzed as described in "Experiment Procedures". The data are expressed as percent distribution (mean ± SD of triplicate determinations) of radioactivity among the lipids.

Fig. 2c Effect of F68 on ^{14}C-oleic acid uptake by phospholipids



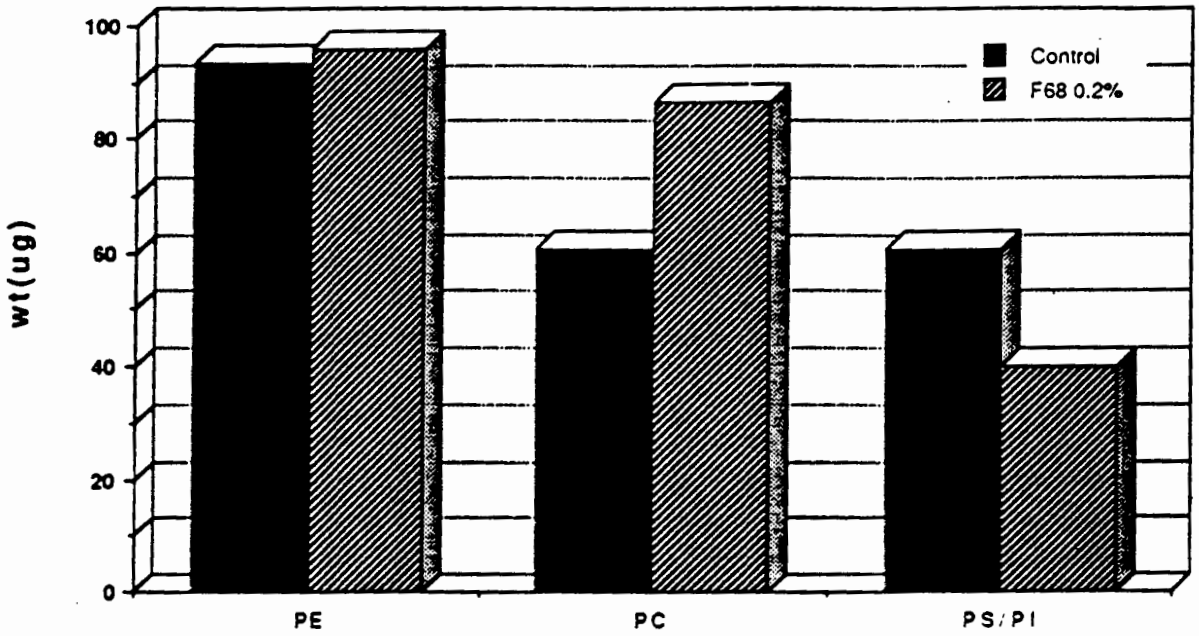
Sf9 cells (1 x 10^6) were incubated with [^{14}C]oleic acid (0.1 μCi/dish) as described in "Experimental Procedures". After labeling, cells were washed twice with HBS and incubated in 1 ml TNM-FH medium containing 10% FBS in the presence and absence of 0.2% (w/v) F68 at the times specified. The reaction was terminated by adding 4 ml of chloroform-methanol (2:1, v/v). Cellular lipid levels were analyzed as described in "Experiment Procedures". Results are mean of triplicate determinations.) are representative of three experiments.

Fig. 3 Effect of serum on ^{14}C-oleic acid uptake



Sf9 cells (1 x 10^7) were grown in TNM-FH medium containing 10% fetal bovine serum (FBS). After 4 days of incubation, cells were harvested by trypsin and resuspended in 1 ml HBS. Cellular lipids were extracted by adding 4 ml of chloroform/methanol (2:1, v/v). Cellular lipid levels were analyzed as described in "Experiment Procedures". The data are expressed as percent distribution of the fatty acids components in total phospholipids. Results are mean ± SD of triplicate determinations.

Fig. 4a Effect of F68 on phospholipids



Sf9 cells (1×10^7) were grown in TNM-FH medium containing 2% fetal bovine serum (FBS) and 0.2% (w/v) F68. After exposure F68 to 4 days, cells were harvested by trypsinization and resuspended in 1 ml HBS. Cellular lipids were extracted and analyzed by GLC as described in "Experiment Procedures". The data are expressed as percent distribution of fatty acids in individual phospholipids. Results are representative of repeated experiments.

Fig. 4b Effect of F68 on fatty acid composition in PE

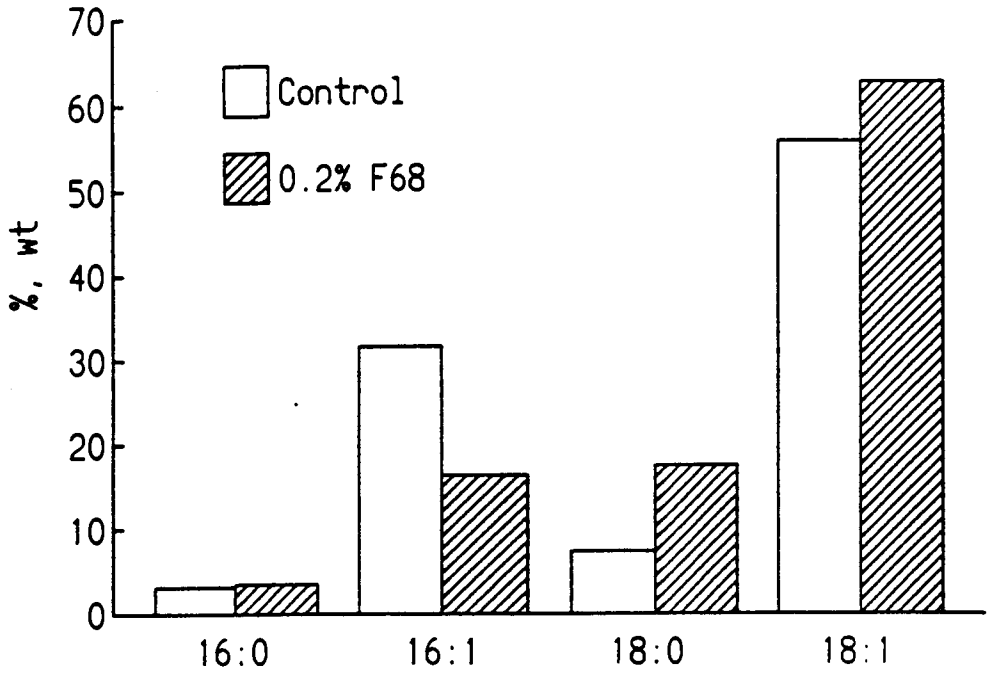
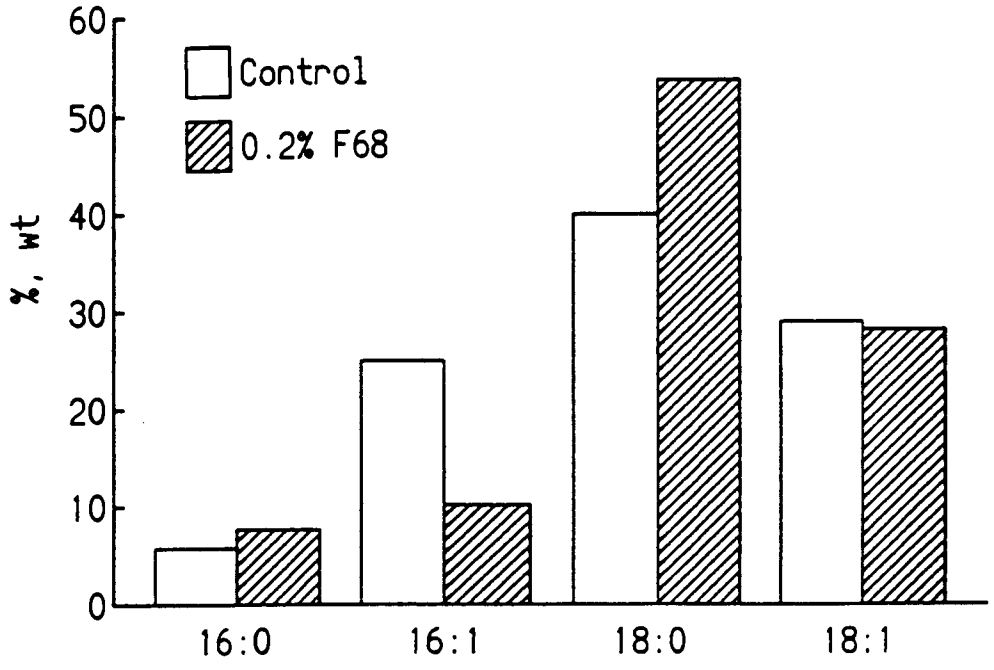


Fig. 4c Effect of F68 on fatty acid composition in PS/PI



Sf9 cells (1 x 10^7) were grown in TNM-FH medium containing 2% fetal bovine serum (FBS) and 0.2% (w/v) F68. After exposure F68 to 4 days, cells were harvested by trypsinization and resuspended in 1 ml HBS. Cellular lipids were extracted and analyzed by GLC as described in "Experiment Procedures". The data are expressed as percent distribution of fatty acids in individual phospholipids. Results are representative of repeated experiments.

A NOVEL APPROACH TO UNDERSTANDING THE ANTIMICROBIAL ACTIVITY OF PEPTIDES

Naveen Pathak, Marie-Helene Janna, Gael Ruche, David McCarthy, and Roger Harrison

School of Chemical Engineering and Material Science, University of Oklahoma

ABSTRACT

A multilinear regression model has been developed to determine the dependence of the antimicrobial activity upon four parameters: hydrophobic moment, hydrophobicity, α-helix fraction, and β-sheet fraction. These parameters were calculated or measured for eight antimicrobial peptides. Antimicrobial activity was quantified as the reciprocal of the measured minimum inhibitory concentration (MIC) against *E. coli*. The regression model revealed that hydrophobic moment had by far the most influence on the antimicrobial activity. The model may be useful in designing new, more active peptides.

INTRODUCTION

Antimicrobial peptides are widely distributed in nature and play an important role in the host defence mechanism of mammals, amphibians, and insects. These peptides are small in size and are therefore easy to modify. The requirement for a rational modification is a mathematical model based on the structure and the properties which can explain their activity.

Magainin 2 (see Fig. 1) is one such peptide derived from the skin of *Xenopus leavius* [1]. Magainin 2 has been shown to form anion selective ion channels that span the microbial membrane [2]. This kind of multimeric aggregation is possible because magainin 2 has a highly amphiphilic nature and a high potential for helix formation according to the Chou and Fasman predictions [3]. The thought of possible antimicrobial activity of the antisense peptide to magainin 2 is based on the complementary relationship between the amino acid assignment of the sense and the antisense codons. The genetic code is biased such that the codon for a hydrophobic amino acid will have an antisense codon coding for a hydrophilic amino acid and vice versa [4]. This relationship will cause an antisense peptide to be structurally inverted with respect to magainin 2, thus conserving the amphiphilic nature. On the basis of this idea several analogs were designed by substitution of amino acids at specific sites of the antisense peptide and tested for activity.

In this work we present a mathematical model which correlates the activity of the peptides to the secondary structure, hydrophobicity, and the amphiphilicity of the peptides. The activity is also analysed on the basis of the hydrophobic moment plot.

MATERIALS AND METHODS

Synthesis of the peptides. The peptides were synthesized by Molecular Biology Resource Facility at the University of Oklahoma Health Science Center. The amino acid composition of each peptide was verified by complete hydrolysis followed by cation exchange chromatography.

Determination of the activity. The *in vitro* antimicrobial activities of the peptides were tested by the microdilution technique [5]. The medium used was tryptic soy broth (TSB). In each well of the Corning Multiple Well Tissue Culture Plate, 100 µl of double concentrated peptide was added to 100 µl of the medium at 10^5 CFU/ml. After incubation for 18 hr at 37^0 C, MIC was the minimal concentration of the peptide for which there was no detectable growth. For each clear well, 10 µl were spotted on an LB 1.5 g/l agar plate and incubated for 24 hr. The MBC was the minimal concentration for which 99.9% of the microorganisms were killed. Each test was done in duplicate, and the results of this test was reconfirmed on a seperate day.

Determination of the secondary structure. Secondary structures of the peptides were determined using circular dichroism (CD) spectroscopy. The CD spectra was recorded on an AVIV Circular Dichroism Spectrophotometer (Model 62 DS) using a quartz cell of 10 mm path length. Solutions used were prepared by mixing 25 mM pottasium phosphate buffer (pH adjusted to 7) and trifluoroethanol in 60:40 ratio. The concentration of the peptides in their final solutions was 30 µg/ml. The spectra was analyzed by PROSEC software to determine the secondary structure.

The **hydrophobicity** of the peptides was quantified by the mean hydrophobicity [6] which is given by:

$$\langle H \rangle = \frac{\sum_{n=1}^{23} H_n}{23}$$

Where H_n is the hydrophobicity of the nth residue as given by the normalised consensus scale of Eisenberg [6].

The magnitude of **amphiphilicity** of the peptides was measured by the mean hydrophobic moment [6] which is equal to

$$\mu_H = \frac{\left\{\left[\sum_{n=1}^{23} H_n \sin(\delta n)\right]^2 + \left[\sum_{n=1}^{23} H_n \cos(\delta n)\right]^2\right\}^{\frac{1}{2}}}{23}$$

For an idealized alpha helix $\delta = 100$.

Multilinear regression was carried out using the Advanced Math facility of Quatro Pro software.

The **hydrophobic moment plot** [6] was constructed by running a window of 11 along the peptide sequence and then selecting the window with highest mean hydrophobic moment. The mean hydrophobicity and the mean hydrophobic moment of this window were plotted. The limits of the plots and the empirical boundaries separating surface, transmembrane, and globular regions were the same as used by Eisenberg [6].

RESULTS AND DISCUSSION

Construction of the antisense peptide and its analogs. The sequence of the antisense peptide was derived from the cDNA sequence that was reported for magainin 2 [1]. Three different types of analogs, types A, B, and C were designed (Fig.2). Replacement in the 15th position was made in order to replace cysteine as it is prone to oxidation. In type A cysteine was replaced by alanine as it is the most similar to cysteine in size, and in type B and C it was replaced by lysine in order to add positive charge on the hydrophilic face. The replacement of glutamine with glycine was made to study the effect of using a residue that tends to break an α-helix (glycine) instead of one that has good probability of being in an α-helix (glutamine).

Antimicrobial activity of all the peptides against *Eschirichia coli, Proteus mirabilis, Klebsiella pneumoniae, Staphylococcus aureus* and *Candida albicans* are shown in the Table I. In the case of the antisense peptide it was important to carry out the experiments in the presense of 1 mM DDT as a reducing agent to avoid oxidation of the cysteines. This DDT concentration did not affect the activity of the antisense peptide in any other way as the MIC value of DDT was found to be 20 mM. Against *E. coli* type C peptides turned out to very active having MIC of 20 μg/ml (same as magainin 2). Type B peptides had moderate activity whereas the activity of type A was quite low against *E. coli.* The peptides had a very low activity against other microorganisms. This might be due to the fact that the length of these peptides might not be sufficient to penetrate the thickness of the lipid bilayer in these microorganisms.

Secondary structure of the peptides. The secondary structure determined by the circular dichroism experiment showed some correlation with the activity against *E. coli.* A2 had the lowest helix content and had lowest activity. Analogs C1, C2, and B1 which had a higher helix content also had higher activity but no definite trend was observed. As the experiments were carried out in 40% trifluoroethanol (which was used to simulate the hydrophobic environment of the membrane), these values may not reflect the exact secondary structure the peptides might attain in the lipid bilayer. These values can be compared with each other to see the relative probability of peptides to form a secondary structure.

In the calculation of the hydrophobic moment the peptides were assumed to be ideal α-helices. This was done due to the lack of actual secondary structure data in the membrane. A strong relationship was observed between the activity and the hydrophobic moment (Table III).

Regression analysis. A multilinear regression of the activity with the mean hydrophobicity, mean hydrophobic moment, and the secondary structure (helix and sheet) revealed a very good fit of the data. The coefficient of determination was found to be 0.96. The correlation obtained is:

$$1/MIC = .085 \langle H \rangle + .341 \langle M \rangle + -0.012 \langle Helix \rangle -0.004 \langle Sheet \rangle -.116$$

Figure 3 shows the goodness of fit between the actual and predicted data.

This correlation was normalised by dividing each variable by its highest value. This was done to compare the influence of all the variables on the activity of the peptides. The normalised equation is:

$$1/MIC = .0147 \langle \overline{H} \rangle + .176 \langle \overline{M} \rangle + -0.005 \langle \overline{Helix} \rangle -0.0008 \langle \overline{Sheet} \rangle -.116$$

where the variables now (with the bar) represent the normalised values.

The coefficient of the mean hydrophobic moment in the normalised equation is much larger than that of any other variable showing that the mean hydrophobic moment, i.e. the amphiphilicity of the peptides, is by far the most important factor influencing the activity of the peptides.

On the hydrophobic moment plot (see Fig. 4) all the peptides lie in the surface region. It is possible that the peptides are surface seeking and aggregate on the surface to form a channel like structure which then penetrates the membrane. This surface seeking tendency of the peptides agrees with the results obtained for some other cytotoxic peptides [6]. An increase in the activity of the peptide was observed with a decrease in the mean hydrophobicity of the chosen window of 11 residues.

CONCLUSIONS

The result of the regression analysis showed that the hydrophobic moment, i.e the amphiphilicity of the peptide, is by far the most important factor influencing the activity of the peptides. The affect of secondary structure on the activity could not be gauged accurately as the experiments for their determination were performed in a medium which only simulated the lipid bilayer. All the peptides turned out to be surface seeking according to the hydrophobic moment plot analysis. These leads to the possibility that the aggregation of the peptides to form ion channels might be taking place before their entering the membrane. A specific trend was also observed between the activity and the position of the peptides on the hydrophobic moment plot. These results can be used to design peptides with higher antimicrobial activity.

REFERENCES

[1] Zasloff, M. (1987) Proc. Natl. Acads. Sci USA, 5449-5453.

[2] Juretic, D., Chen, H.C, Brown, J.H., Morell, J.L.,Hendler, R.W., and Westerhoff, H.V. (1989) FEBS Lett. 249,219-223.

[3] Chen, H.C., Brown, J.H., Morell, J.L., and Huang, C.M. (1988) FEBS Lett. 236, 462-466.

[4] Blalock, J.E, and Smith, E.M. (1984) Biochem. Biophys. Res. Comm. 121, 203-207.

[5] Jones, R.N., Barry, A.L., Gavan, T.L., Washington, J.A. (1985) In Lennette, E.H., Balows, A., Hausler, W.J., Shadomy, H.J. eds. Manual of Clinical Microbiology, 4th ed. (American Society for Microbiology, Washington, D.C.).

[6] Eisenberg, D., Schwartz, E., Komaromy, M., Wall, R., (1984) . J. Mol. Biol. 179: 125-142.

Table I. Antimicrobial activities of antisense peptide of magainin antisense peptide analogs compared to that of magainin 2.

organism (ATCC no.)	MIC/MBC (μg/ml)[a]							
	Magainin 2	Antisense peptide + DTT[b]	Analog type A		Analog type B		Analog type C	
			A1	A2	B1	B2	C1	C2
E.coli D31	20/40	60/80	100/100	400/>500[c]	40/40	40/60	20/40	20/20
K.pneumo. (13883)	80/100	>500/>500	>500/>500	>500/>500	300/500	500/500	300/>500	300/300
C.albicans (10231)	400/400	>500/>500	>500/>500	>500/>500	300/400	>500/>500	300/400	>500/>500
P.mirabilis (14153)	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500
S.aureus (6538)	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500	>500/>500

[a] MIC = minimal inhibitory concentration, MBC = minimal bactericidal concentration

[b] Dithiothreitol (1 mM)

[c] The symbol ">" preceeding a peptide concentration indicates that no effects on survival were detected over a range of concentrations including the concentration shown.

TABLE II

Secondary structure estimated from circular dichroism measurements in 40% trifluoroethanol

Peptide	Calculated %			
	α-Helix	ß-Sheet	Turn	Random
Antisense peptide	27	0	30	43
Analog A1	42	8	35	15
Analog A2	21	11	33	35
Analog B1	44	0	41	15
Analog B2	30	0	37	33
Analog C1	35	0	37	28
Analog C2	38	0	37	25
Magainin 2	22	20	24	34

TABLE III

Data used in the multiple linear regression analysis

Peptide	Mean Hydrophobicity	Mean Hydrophobic Moment	Fraction α-Helix	Fraction ß-Sheet	MIC*	1/MIC
Antisense Peptide	0.090	0.367	0.27	0	60	0.0167
Analog A1	0.105	0.352	0.42	0.08	100	0.0100
Analog A2	0.160	0.325	0.21	0.11	400	0.0025
Analog B1	0.013	0.444	0.44	0	40	0.0250
Analog B2	0.070	0.417	0.30	0	40	0.0250
Analog C1	-0.060	0.518	0.35	0	20	0.0500
Analog C2	-0.006	0.487	0.38	0	20	0.0500
Magainin 2	0.173	0.455	0.22	0.20	20	0.0500

* MIC = minimal inhibitory concentration

Magainin 2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
	H2N - Gly	Ile	Gly	Lys	Phe	Leu	His	Ser	Ala	Lys	Lys	Phe	Gly	Lys	Ala	Phe	Val	Gly	Glu	Ile	Met	Asn	Ser - COOH	
5'	GGA	ATT	GGT	AAA	TTT	TTG	CAC	TCA	GCA	AAA	AAA	TTT	GGA	AAA	GCT	TTT	GTG	GGA	GAG	ATA	ATG	AAT	TCA	3'
3'	CCT	TAA	CCA	TTT	AAA	AAC	GTG	AGT	CGT	TTT	TTT	AAA	CCT	TTT	CGA	AAA	CAC	CCT	CTC	TAT	TAC	TTA	AGT	5'
	HOOC - Ser	Asn	Thr	Phe	Lys	Gln	Val	AGG Gly	Cys	Phe	Phe	Lys	Ser	Phe	Ser	Lys	His	Ser	Leu	Tyr	His	Ile	AGG Gly - NH2	
	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	

Antisense peptide

Figure 1: Amino acid sequence of magainin 2 and its antisense peptide.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
G	I	G	K	F	L	H	S	A	K	K	F	G	K	A	F	V	G	E	I	M	N	S	Magainin 2
G	I	H	Y	L	S	H	K	S	F	S	K	F	F	C	G	V	Q	K	F	T	N	S	Antisense peptide
G	I	H	Y	L	S	H	K	S	F	S	K	F	F	A	G	V	Q	K	F	T	N	S	Analog A1
G	I	H	Y	L	S	H	K	S	F	S	K	F	F	A	G	V	G	K	F	T	N	S	Analog A2
G	I	H	Y	L	S	H	K	S	F	S	K	F	F	K	G	V	Q	K	F	T	N	S	Analog B1
G	I	H	Y	L	S	H	K	S	F	S	K	F	F	K	G	V	G	K	F	T	N	S	Analog B2
G	I	H	K	L	S	H	K	S	F	S	K	F	F	K	G	V	Q	K	F	T	N	S	Analog C1
G	I	H	K	L	S	H	K	S	F	S	K	F	F	K	G	V	G	K	F	T	N	S	Analog C2

Figure 2: Amino acid sequence of the various analogs showing the substitutions.

Figure 3

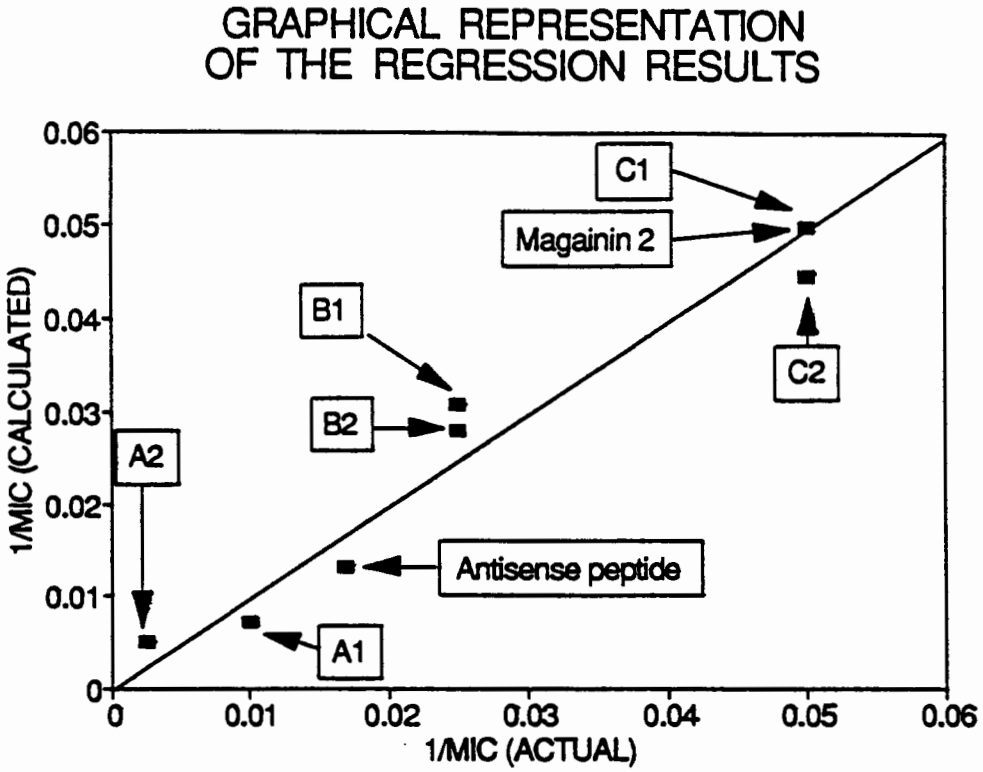
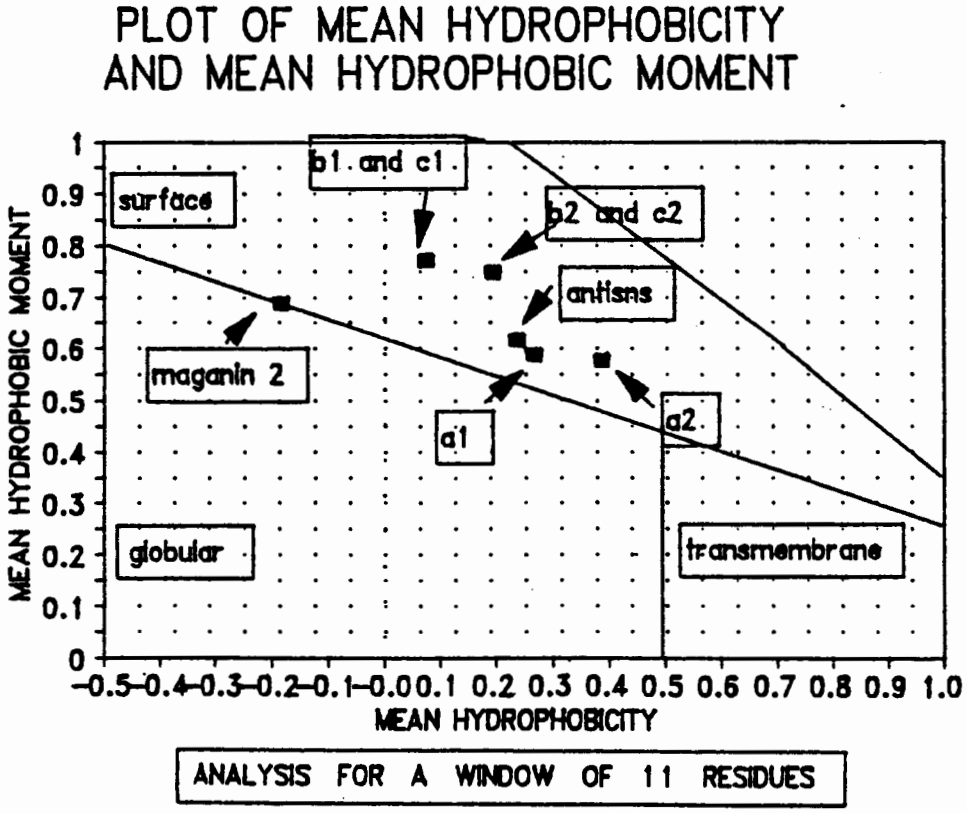


Figure 4



MASS TRANSFER IN THE BIOREMEDIATION OF SOILS CONTAMINATED WITH TRAPPED NON-AQUEOUS PHASE LIQUIDS

Xiaoqing Yang, Larry E. Erickson, and L.T. Fan
Department of Chemical Engineering, Durland Hall
Kansas State University, Manhattan, KS 66506

ABSTRACT

A model has been developed to study the mass transfer of trapped non-aqueous phase liquids (NAPL) in saturated soils. This model considers the immobile NAPL phase as discrete blobs while the aqueous phase as a continuum. The blobs of the NAPL may contain several mutual miscible components and have a global mass distribution over the soil bed as well as a local size distribution. Interactions between the two liquid phases are incorporated into the governing equations for the aqueous phase. The rates of dissolution and desorption are assumed to be of the first order. Simplified simulations have been carried out to examine the characteristics of the system, such as time-dependent concentration profiles, remaining quantities of the contaminant, and variations of the NAPL blob size.

Key words: NAPL, transport, soil, mass transfer.

INTRODUCTION

The widespread release of organic chemicals in the environment frequently leads to soil contamination with non-aqueous phase liquids (NAPL) because many of these compounds are barely soluble in water (Williams and Wilder, 1971; Lyman et al., 1990). Understanding the mechanism of mass transfer of organics in subsurface soils is essential for design and implementation of bioremediation technology.

To date, multiphase transport models in soils have mainly been developed by extending Darcy's law. In those models, each fluid phase is naturally assumed to be a continuum so that the classical laws of continuum mechanics apply (see, e.g., Peaceman, 1977; Kueper and Frind, 1991). Mass transfer between phases is frequently described by a local equilibrium assumption (see, e.g., Abriola and Pinder, 1985; Baehr and Corapcioglu, 1987). According to field observations and experimental results, NAPL in groundwater may be trapped as blobs, i.e., completely surrounded by water and solid materials. Because of the interfacial tension and capillary effect, NAPL blobs are often immobile and stable under aqueous flow conditions. Mass transfer between phases depends on concentration difference, interfacial area and local hydrodynamic environment (Feenstra and Coburn, 1986; Geller and Hunt, 1993).

The objective of this research is to develop a model and simulate numerically the dissolution and transport of organic contaminants in saturated subsurface soils. In the model, the NAPL is treated as discrete blobs, and mass transfer between phases is characterized by first order kinetics.

MODEL DEVELOPMENT

Pore space in the soil bed under consideration is regarded as being completely occupied by NAPL blobs and/or aqueous solution; therefore, this gives rise to a three-phase system. Adsorption-desorption occurs between the solid phase and liquid phases. The NAPL may contain several mutually soluble components. The NAPL blobs may have a number distribution $\lambda(\underline{x})$ over the soil bed and a local size distribution $f(\underline{x},t,v)$. Compounds in the NAPL phase dissolve into the aqueous phase where convective-dispersive solute transport occurs.

The NAPL phase is treated as discrete blobs which do not undergo flow, coalescence, and breakage; the aqueous phase is a continuum. Contaminants in the solid phase are considered to be adsorbed on the surface of the particles. Under these considerations, the governing equations have been derived from the mass conservation laws as follows:

variation of the NAPL volume fraction

$$\varepsilon^{\alpha}=\int_0^{\infty}\lambda(\underline{x})\,f(\underline{x},t,v)\,v(\underline{x},t)\,dv \tag{1}$$

NAPL dissolution

$$\frac{\partial(\rho^{\alpha}\omega_i^{\alpha}v)}{\partial t}=-\gamma\sigma v k_{is}\left(\frac{\omega_i^{\alpha}}{K_{inl}}-\frac{q_i^{s}}{K_{ils}}\right) - (1-\gamma)\,\sigma v k_{in}\left(\frac{\omega_i^{\alpha}}{K_{inl}}-C_i^{\beta}\right) \tag{2}$$

aqueous phase solute transport

$$\frac{\partial(\varepsilon^{\beta}C_i^{\beta})}{\partial t}=\nabla\cdot(\varepsilon^{\beta}E_i^{\beta}\nabla C_i^{\beta})-\nabla\cdot(\varepsilon^{\beta}\underline{U}^{\beta}C_i^{\beta}) + [a-\int_0^{\infty}f\lambda\gamma\sigma v dv]\,k_{il}\left(\frac{q_i^{s}}{K_{ils}}-C_i^{\beta}\right) + \int_0^{\infty}f\lambda\,(1-\gamma)\,\sigma v k_{in}\left(\frac{\omega_i^{\alpha}}{K_{inl}}-C_i^{\beta}\right)dv \tag{3}$$

solid phase mass balance

$$\frac{\rho^s \partial[(1-\varepsilon) q_i^s]}{\partial t} = -[a - \int_0^\infty f\lambda\gamma\sigma v dv] k_{il} (\frac{q_i^s}{K_{ils}} - C_i^\beta) + \int_0^\infty f\lambda\gamma\sigma v k_{is} (\frac{\omega_i^\alpha}{K_{inl}} - \frac{q_i^s}{K_{ils}}) dv \quad (4)$$

These governing equations are subjected to the following natural constaints;

blob size distribution constraint

$$\int_0^\infty f(x, t, v) dv = 1 \quad (5)$$

$$f(x, t_1, v_1(x, t_1)) = f(x, t_0, v_0(x, t_0)) \quad (6)$$

NAPL composition constraint

$$\sum_i \omega_i^\alpha(x, t, v) = 1 \quad (7)$$

In deriving the governing equations, local equilibrium is not assumed; instead, the mass transfer between phases is assumed to obey the first order kinetics. With the appropriate initial and boundary conditions, these equations constitute a mathematically closed model.

CASE STUDY

Numerical simulations have been carried out for a problem illustrated in Figure 1. The simplifying assumptions include the following.

1. One-dimensional flow and transport prevail in the soil bed.
2. The NAPL phase contains only one component. Initially, the NAPL blobs have unique size and have a constant number distribution over the part of the bed: $0<x<L_1$.
3. The NAPL is a perfectly nonwetting fluid, i.e., the solid contacting area of the NAPL blobs is zero.
4. The NAPL blobs are spheres.

With these assumptions, the governing equations within the NAPL contaminated part of the soil bed can be easily simplified as follows:

variation of the NAPL volume fraction

$$\varepsilon^\alpha = \frac{4}{3}\pi R^3 \lambda \quad (8)$$

NAPL dissolution

$$\rho^{\alpha}\frac{\partial R}{\partial t}=-k_n(C_{sat}-C^{\beta}) \tag{9}$$

aqueous phase solute transport

$$\frac{\partial(\varepsilon^{\beta}C^{\beta})}{\partial t}=\frac{\partial}{\partial x}\left[E\frac{\partial(\varepsilon^{\beta}C^{\beta})}{\partial x}\right]-\frac{\partial(\varepsilon^{\beta}uC^{\beta})}{\partial x} +k_1a\left(\frac{q^s}{K_{1s}}-C^{\beta}\right)+4\pi\lambda R^2k_n(C_{sat}-C^{\beta}) \tag{10}$$

solid phase mass balance

$$\rho^{s}(1-\varepsilon)\frac{\partial q^s}{\partial t}=-ak_1\left(\frac{q^s}{K_{1s}}-C^{\beta}\right) \tag{11}$$

The governing equations within the remaining part of the soil bed, $L_1<x<L$, are as follows:

aqueous phase solute transport

$$\frac{\partial(\varepsilon^{\beta}C^{\beta})}{\partial t}=\frac{\partial}{\partial x}\left[E\frac{\partial(\varepsilon^{\beta}C^{\beta})}{\partial x}\right]-\frac{\partial(\varepsilon^{\beta}uC^{\beta})}{\partial x}+k_1a\left(\frac{q^s}{K_{1s}}-C^{\beta}\right) \tag{12}$$

solid phase mass balance

$$\rho^{s}(1-\varepsilon)\frac{\partial q^s}{\partial t}=-ak_1\left(\frac{q^s}{K_{1s}}-C^{\beta}\right) \tag{13}$$

The porosity, ε, and the volumetric flow rate, Q_0, are constant; moreover,

$$\varepsilon^{\alpha}+\varepsilon^{\beta}=\varepsilon \tag{14}$$

$$\varepsilon^{\beta}u=Q_0=\varepsilon u_0 \tag{15}$$

The boundary conditions have been derived from the continuity of concentration profiles and that of mass flux, thereby yielding

$$0=C^{\beta}\big|_{(x=0^{+},\,t=t)}-\frac{E}{u}\frac{\partial C^{\beta}}{\partial x}\Big|_{(x=0^{+},\,t=t)} \tag{16}$$

$$C^{\beta}\big|_{x=L_1^{-}}=C^{\beta}\big|_{x=L_1^{+}} \tag{17}$$

$$\frac{\partial C^{\beta}}{\partial x}\Big|_{x=L_1^-} = \frac{\partial C^{\beta}}{\partial x}\Big|_{x=L_1^+} \qquad (18)$$

$$\frac{\partial C^{\beta}}{\partial x}\Big|_{(x=L,\ t=t)} = 0 \qquad (19)$$

In simulating the system, the governing equations in the form of coupled non-linear partial differential equations have been solved by a finite difference scheme (Yang et al., 1992). The parameter values used in the simulations are listed in Table 1.

Table 1. Parameter values for the simulations

$a=10^2 cm^2/cm^3$	C_{sol}=500ppm	$E=1.157\times10^{-5} cm^2/s$	K_{ls}=1.0L/kg	$k_l=0.579\times10^{-7} cm/s$
k_n=0.173m/d	L_1=0.33m	L=1.0m	R_0=1.0mm	$\lambda_0=0.6\times10^{+7}/m^3$
u_0=0.2m/d	ε=0.5	$\rho^{\alpha}=1.0 g/cm^3$	ε_0^{α}=0.05	$\rho^{s}=2.5 g/cm^3$

RESULTS AND DISCUSSION

The initial contaminant distributions are sketched in Figure 2. As given in Table 1, the NAPL volume fraction, ε_0^{α}, is 0.05; since the porosity of the soil bed, ε, is 0.5, the NAPL phase occupies 10% of the void volume. The major part of the contaminant, however, is in the NAPL phase. The total amount is the sum of the contaminants in the NAPL phase, the adsorbed phase, and the aqueous phase.

The total contaminant distributions at various times are plotted in Figure 3, which indicates the rates and degrees of contaminant removal at various locations in the soil bed as functions of time. The contaminant distribution at 100 days is presented in Figure 4. Figure 5 illustrates dissolution and desorption at a specific location near the front end of the soil bed. It reveals that the desorption and dissolution occur simultaneously; however, the dissolution is completed in 34 days, but the desorption continues after 60 days. Figure 6 presents the NAPL distribution and dimensionless aqueous concentration profiles at various times. At any time, NAPL dissolution takes place over a narrow band which is about 10 cm wide. In this case, the groundwater flow velocity is relatively small (0.2m/d); therefore, transport is under convection-dispersion control. Under such a condition, the downstream solution is saturated as long as sufficient NAPL is present.

CONCLUSIONS

A discrete blob model governing the mass transfer in saturated soils contaminated with immobile non-aqueous phase liquids has been proposed. A case study indicates that dissolution and desorption occur simultaneously; however, the dissolution is completed first. The NAPL phase attenuates via a moving front, and under convection-dispersion control conditions, the downstream solution is saturated as long as sufficient NAPL is present.

ACKNOWLEDGMENT

Although the research described in this article has been funded in part by the United States Environmental Protection Agency under assistance agreement R-815709 to the Great Plains-Rocky Mountain Hazardous Substance Research Center for U.S. EPA Region 7 and 8 with headquarters at Kansas State University, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the agency and no official endorsement should be inferred. This research was partially supported by the Kansas State University Center for Hazardous Substance Research.

NOMENCLATURE

a	interfacial area of solid phase per unit volume of the soil bed, L^2/L^3
C	concentration in the aqueous phase, M/L^3
C_{sol}	solubility of the substance in the β phase, M/L^3
E	dispersive coefficient, L^2/t
f	distribution density of NAPL blob size
k_{il}, k_l	mass transfer coefficient, L/t
K_{ils}	adsorption partition coefficient, L^3/M
k_{in}, k_n	mass transfer coefficient, L/t
K_{inl}	NAPL partition coefficient, L^3/M
k_{is}	mass transfer coefficient, L/t
L	length of the soil bed, L
L_1	length of the part of the soil bed contaminated with the organic phase, L
q	concentration in the solid phase, M/M
R	radius of the NAPL blob, L
t	time, t
u	flow velocity, L/t
u_0	flow velocity in the part of the soil bed without the NAPL phase, L/t
v	volume of a NAPL blob, L^3
x	axis, L
X	relative distance along the soil bed

Greek letters

λ number distribution of NAPL blobs, $1/L^3$
σ surface area-volume ratio of the NAPL blob, 1/L
γ ratio of solid contacting area to the surface area of the NAPL blob
ρ density, M/L^3
ε void fraction of the soil bed
ε^{α} NAPL phase volume fraction
ε^{β} aqueous phase volume fraction
ω_i^{α} weight fraction of species i in the NAPL phase

superscript

α NAPL phase, β aqueous phase, s solid phase

subscript

i species i, $_0$ initial value

underline vector

REFERENCES

Abriola, L.M. and G.F. Pinder, A multiphase approach to the modeling of porous media contamination by organic compounds, 2. Numerical simulation, *Water Resour. Res.*, **21**, 19-26(1985).

Baehr, A.L., and M.Y. Corapcioglu, A compositional multiphase model for groundwater contamination by petroleum products, 2. numerical solution, *Water Resour. Res.*, **23**, 201-213(1987).

Feenstra, S. and J. Coburn, Subsurface contamination from spills of denser than water chlorinated solvents, *Calif. WPCF Bull.*, **23**(4), 26-34(1986).

Geller, J.T. and J.R. Hunt, Mass transfer from nonaqueous phase organic liquids in water-saturated porous media, *Water Resour. Res.*, **29**, 833-845(1993).

Kueper, B.H., and E.O. Frind, Two-phase flow in heterogeneous porous media: 1. model development, *Water Resour. Res.*, **27**, 1049-1057(1991).

Lyman, W. J., D.C.Noonan, and P.J. Reidy, Cleanup of Petroleum Contaminated Soils at Underground Storage Tanks, pp.3-216, Noyes Data, Park Ridge, NJ, 1990.

Peaceman, D.W., Fundamentals of Numerical Reservoir Simulation, Elsevier Science, New York, 1977.

Williams, D.E., and D.G. Wilder, Gasoline pollution of a groundwater reservoir-- a case history, *Ground Water,* **9**(6), 50-54(1971).

Yang, X., L.E. Erickson, and L.T. Fan, Dispersive-convective transport in the bioremediation of contaminated clay layers, Proceedings of the Conference of Hazardous Waste Research, L.E. Erickson ed., pp.581-599, Manhattan, KS, 1992.

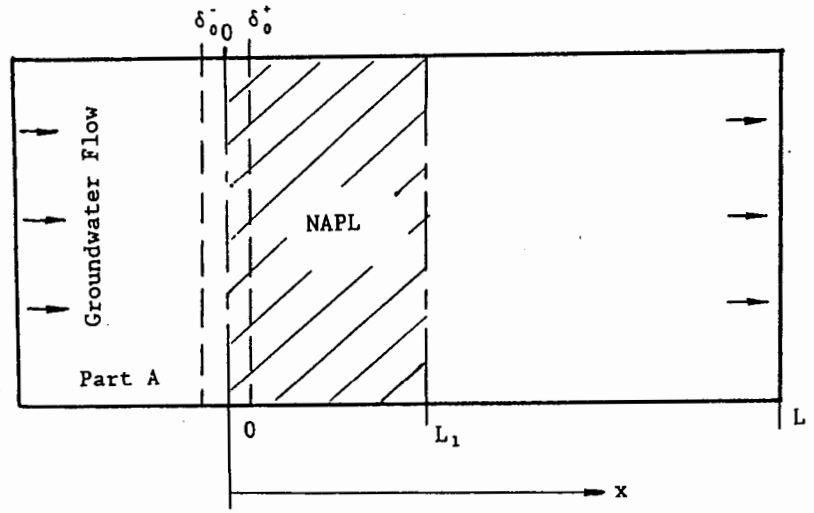


Figure 1. Schematic of the soil bed: part of the bed is contaminated by the NAPL initially.

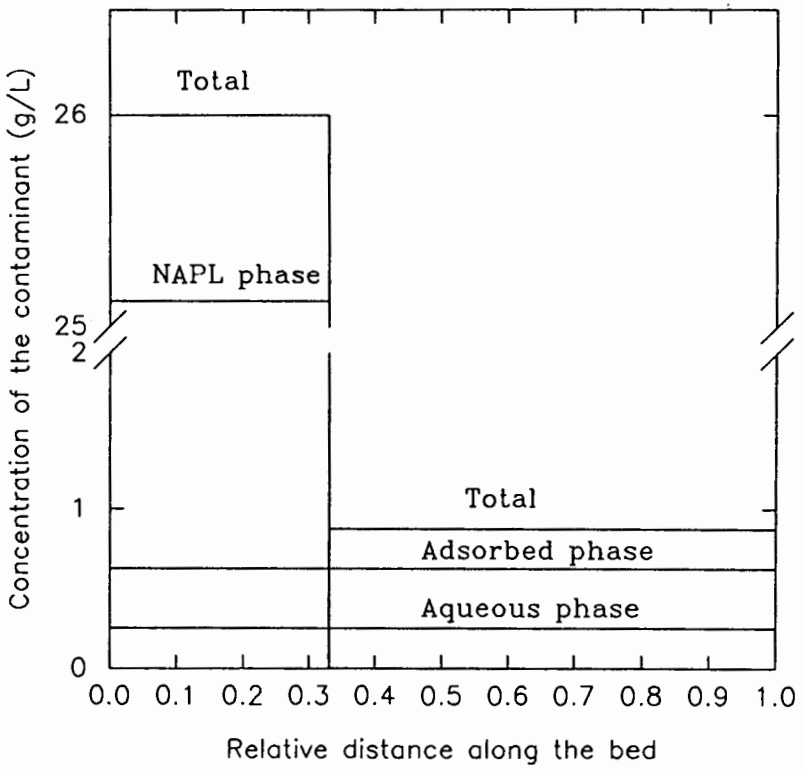


Figure 2. Initial contaminant distribution along the soil bed (g contaminant/L soil bed).

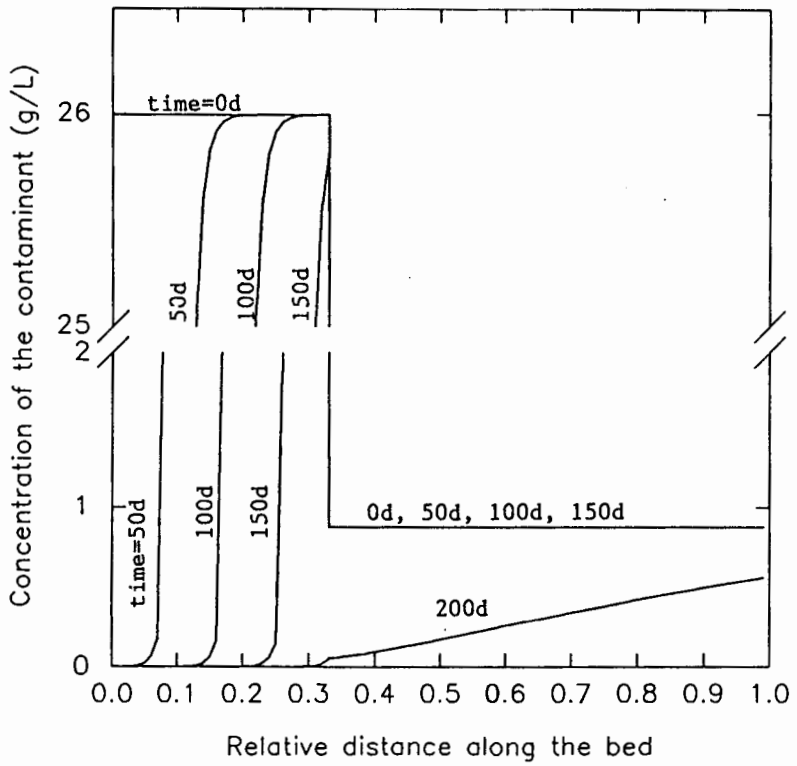


Figure 3. Total contaminant distribution at various times (g contaminant/L soil bed).

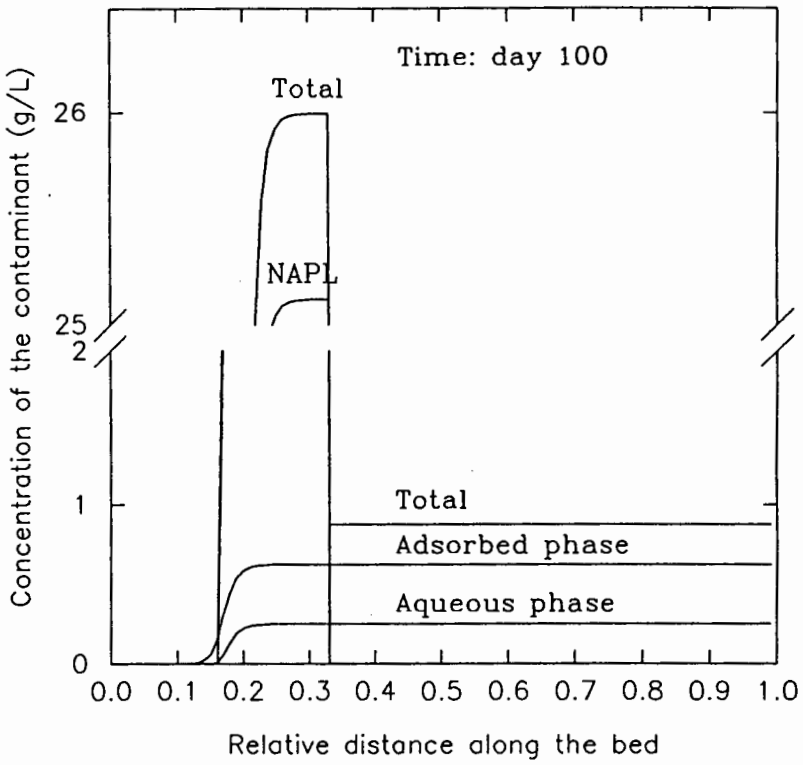


Figure 4. Contaminant distribution after 100 days (g contaminant/L soil bed).

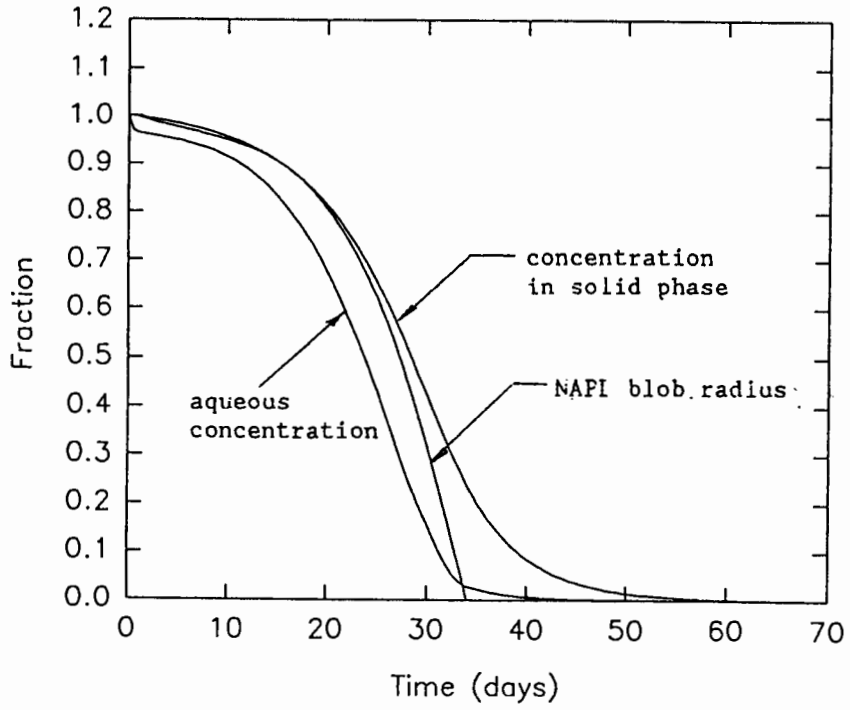


Figure 5. Dissolution and desorption at X=0.04 where NAPL is present initially.

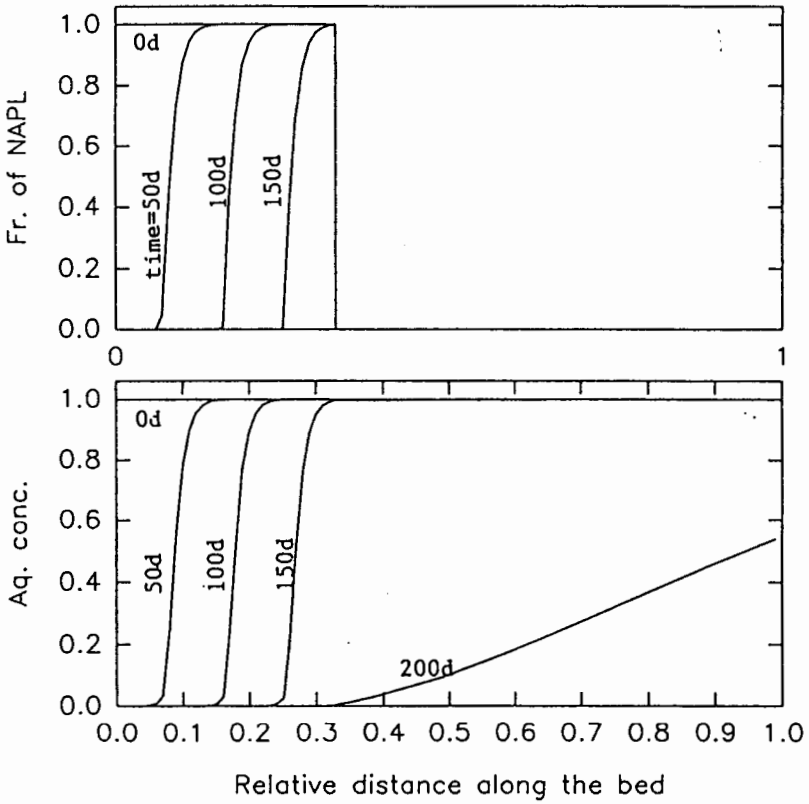


Figure 6. NAPL spatial distributions and dimensionless aqueous concentration profiles at various times.